

133817

STIC-Biotech/ChemLib

From: Schnizer, Holly
Sent: Tuesday, September 28, 2004 8:43 PM
To: STIC-Biotech/ChemLib
Subject: seq. search for appl. no. 10/049,399

Please search all databases for the following sequence:

residues 2169-2332 of SEQ ID NO:1

Thank you.

Holly Schnizer
AU 1653
Office: REM 3C79
Mailbox: REM 3C70
(571)272-0958

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: 9/29/04
Date Completed: 9/30/04
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence: # 1
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: Q2P
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 16:33:53 ; Search time 55 seconds
(without alignments)
842.505 Million cell updates/sec

Title: US-10-049-399A-1_COPY_2169_2332
Perfect score: 871
Sequence: 1 CDLNSCSPLMGESKAISDA.....VWQIALRMEVLGCEAQDLY 164

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_29Jan04:*
- 1: Geneseq1980s:*
 - 2: Geneseq1990s:*
 - 3: Geneseq2000s:*
 - 4: Geneseq2001s:*
 - 5: Geneseq2002s:*
 - 6: Geneseq2003as:*
 - 7: Geneseq2003bs:*
 - 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	871	100.0	643	6	Abp60514 Human fac
2	871	100.0	684	2	Aar73022 Human fac
3	871	100.0	684	2	Aar74091 Factor-VI
4	871	100.0	790	5	Ades4594 Recombina
5	871	100.0	1014	1	Aap71139 Factor VII
6	871	100.0	1383	2	Aaw33227 Procoagul
7	871	100.0	1383	2	Aaw33228 Procoagul
8	871	100.0	1383	2	Aaw33229 Procoagul
9	871	100.0	1424	1	Aap80268 Modified
10	871	100.0	1424	1	Aap91169 Sequence
11	871	100.0	1424	4	Aab48842 Mutant ma
12	871	100.0	1424	5	Aao18622 Human mat
13	871	100.0	1425	1	Aap80267 Modified
14	871	100.0	1438	3	Aab01262 B-domain
15	871	100.0	1440	3	Aar12971 Factor VI
16	871	100.0	1445	5	Abg92540 LE B-doma
17	871	100.0	1447	5	Abg92541 SArg B-do
18	871	100.0	1457	2	Aaw46246 Human fac
19	871	100.0	1457	2	Aaw44372 Human fac
20	871	100.0	1457	2	Aay21675 Beta-doma
21	871	100.0	1459	4	Aae10832 Human fac
22	871	100.0	1471	2	Aaw23414 Human B-d
23	871	100.0	1471	4	Aab67959 Amino aci
24	871	100.0	1516	1	Aap80265 Modified
25	871	100.0	1661	2	Aaw18670 Factor VI

26	871	100.0	2332	1	AAP71726
27	871	100.0	2332	1	AAP71728
28	871	100.0	2332	1	AAP71727
29	871	100.0	2332	1	AAP71729
30	871	100.0	2332	2	AAR43257
31	871	100.0	2332	2	AAR43254
32	871	100.0	2332	2	AAW33225
33	871	100.0	2332	2	AAW33226
34	871	100.0	2332	2	AAW33227
35	871	100.0	2332	2	AAW33223
36	871	100.0	2332	2	AAW44132
37	871	100.0	2332	2	AAW53483
38	871	100.0	2332	2	AAW53484
39	871	100.0	2332	3	AAW57847
40	871	100.0	2332	4	AAE10826
41	871	100.0	2332	4	AAE10826
42	871	100.0	2332	4	AAE10826
43	871	100.0	2332	4	AAE11200
44	871	100.0	2332	5	AAU79869
45	871	100.0	2332	6	ABP60513

ALIGNMENTS

RESULT 1
ABP60514 standard; protein; 643 AA.

- AC ABP60514;
- XX XX
- DT 28-MAR-2003 (first entry)
- DE Human factor VIII light chain P2153Q mutant.
- XX Human; factor VIII; T-cell; haemostatic; gene therapy; light chain;
- KW thrombotic disorder; haemophilia; mutant; mutein.
- XX Homo sapiens.
- XX WO200298454-A2.
- XX 12-DEC-2002.
- XX 31-MAY-2002; 2002WO-BE0000088.
- XX 31-MAY-2001; 2001WO-EP006297.
- PR 14-NOV-2001; 2001GB-00027321.
- XX (COLL-) COLLEN RES FOUND VZW ONDERWIJSEN NAVORSI.
- XX Jacquemin MG, Saint-Remy JR;
- XX WPI; 2003-140553/13.
- DR N-PSDB; ABV99881.

New recombinant modified functional polypeptide, useful as a medicament for treating or preventing hemophilia A or B, comprises multiple point-mutations responsible for T-cell activation reduction or elimination.

Claim 79; Page 73-75; 85pp; English.

The invention relates to a novel recombinant modified functional polypeptide which exerts at least one function of a mammalian protein and which eliminates or reduces by at least about 80%, with respect to activation by the unmodified polypeptide, the activation of at least one T-cell clone derived from a mammal with antibody against the unmodified polypeptide. The polypeptide of the invention has haemostatic activity. The polynucleotide encoding the polypeptide may have a use in gene therapy. The recombinant modified functional polypeptide is useful as a medicament for treating or preventing a thrombotic disorder, such as haemophilia A or B. The peptide is also useful for in vitro or in vivo evaluation and/or modulation of T-cell reactivity. The recombinant point-

CC mutated factor VIII light chain molecule is used as a template for
 CC determining whether modification of the factor VIII molecule is able to
 CC reduce the clearance, e.g. to increase the half-time, of factor VIII in
 CC plasma. The present sequence represents the human factor VIII light chain
 CC P2153Q mutant

XX
 XX
 SQ Sequence 643 AA;

Query Match 100.0%; Score 871; DB 6; Length 643;
 Best Local Similarity 100.0%; Pred. No. 2.9e-90;
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCSMPLGMSKALSDAQITASSYFTNMFTWSPSKARLHLQGRSNARWPQVNNPKE 60
 DB 480 CDLNSCSMPLGMSKALSDAQITASSYFTNMFTWSPSKARLHLQGRSNARWPQVNNPKE 539
 QY 61 WLQVDFQKTMKVTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
 DB 540 WLQVDFQKTMKVTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQD 599
 QY 121 SFTFVNSLDPPLTLRLRIHPQSWVHQIALRMEVLGCEAQDLY 164
 DB 600 SFTFVNSLDPPLTLRLRIHPQSWVHQIALRMEVLGCEAQDLY 643

RESULT 2
 AAR73022
 ID AAR73022 standard; peptide; 684 AA.
 AC AAR73022;
 XX
 XX 25-MAR-2003 (revised)
 DT 21-NOV-1995 (first entry)
 XX
 XX Human Factor-VIII C-terminal fragment.

XX Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.

XX Homo sapiens.

XX WO9513301-A1.

XX 18-MAY-1995.

XX 10-NOV-1994; 94WO-DK000424.

XX 12-NOV-1993; 93DK-00001281.

XX (NOVO) NOVO-NORDISK AS.

XX Persson E;

XX WPI; 1995-194038/25.

XX Crosslinked Factor VIII polypeptide which is stable - is prepnd. using
 PT bis(sulphosuccinimidy)l suberate or disuccinimidyyl suberate in the
 PT presence of polysorbate 80 to produce a coagulant with long lasting
 PT activity.

XX Disclosure; Page 24; 36pp; English.

XX This is the C-terminal fragment (corresponding to AAs 1649-2332) of human
 CC Factor-VIII which may be crosslinked resulting in increased stability and
 CC retention of high activity over extended periods of time after activation
 CC by thrombin. The polypeptide is used to prevent or treat diseases caused
 CC by the absence or deficiency of Factor-VIII in a subject such as
 CC haemophilia. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 684 AA;

Query Match 100.0%; Score 871; DB 2; Length 684;
 Best Local Similarity 100.0%; Pred. No. 3.2e-90;
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCSMPLGMSKALSDAQITASSYFTNMFTWSPSKARLHLQGRSNARWPQVNNPKE 60
 DB 521 CDLNSCSMPLGMSKALSDAQITASSYFTNMFTWSPSKARLHLQGRSNARWPQVNNPKE 580
 QY 61 WLQVDFQKTMKVTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
 DB 581 WLQVDFQKTMKVTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQD 640
 QY 121 SFTFVNSLDPPLTLRLRIHPQSWVHQIALRMEVLGCEAQDLY 164
 DB 641 SFTFVNSLDPPLTLRLRIHPQSWVHQIALRMEVLGCEAQDLY 684

RESULT 3
 AAR74091
 ID AAR74091 standard; protein; 684 AA.
 AC AAR74091;
 XX
 XX 25-MAR-2003 (revised)
 DT 04-NOV-1995 (first entry)
 XX
 XX Factor-VIII light chain C-terminal fragment.

XX human; Factor VIII; light chain; C-terminal fragment; thrombin cleavage;
 KW blood-clotting.

XX Homo sapiens.

XX WO9513300-A1.

XX 18-MAY-1995.

XX 10-NOV-1994; 94WO-DK000423.

XX 12-NOV-1993; 93DK-00001280.

XX (NOVO) NOVO-NORDISK AS.

XX Kjalke M, Ezban Rasmussen M;

XX WPI; 1995-194037/25.

XX Factor VIII polypeptide(s) comprising a heavy chain shorter than native
 PT A1-A2 domain - are easier to produce recombinantly and retain coagulant
 PT activity, may be used to treat patients who have developed antibodies to
 PT C-terminal epitope(s) of Factor VIII.

XX Claim 4; Page 33-35; 51pp; English.

XX The sequence represents C-terminal residues 1649-2332 of a human Factor-
 CC VIII light chain. The sequence is produced by treating a full-length
 CC Factor-VIII polypeptide with a protease, e.g. thrombin. The fragment may
 CC be produced recombinantly in conjunction with a C-terminally truncated
 CC heavy chain to reduce production costs and improve safety, and production
 CC levels and stability are higher than for the full-length form. When the
 CC fragment is used along with a C-terminally truncated heavy chain, it may
 CC be used to treat patients who have developed antibodies against epitopes
 CC in the C-terminal part of the heavy chain. (Updated on 25-MAR-2003 to
 CC correct PN field.)

XX Sequence 684 AA;

Query Match 100.0%; Score 871; DB 2; Length 684;
 Best Local Similarity 100.0%; Pred. No. 3.2e-90;
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCSMPLGMSKALSDAQITASSYFTNMFTWSPSKARLHLQGRSNARWPQVNNPKE 60
 DB 521 CDLNSCSMPLGMSKALSDAQITASSYFTNMFTWSPSKARLHLQGRSNARWPQVNNPKE 580
 QY 61 WLQVDFQKTMKVTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQD 120

Db 581 WLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 640
QY 121 SFTPVVNSLDPPLLRYLRIRHPSQSWHQIALRMEVLGCEAODLY 164
Db 641 SFTPVVNSLDPPLLRYLRIRHPSQSWHQIALRMEVLGCEAODLY 684

RESULT 4
ADE64594
ID ADE64594 standard; protein; 790 AA.
XX AC ADE64594;
XX 29-JAN-2004 (first entry)
XX DE Recombinant blood coagulation factor VIII protein, SEQ ID 4.
XX blood coagulation factor VIII; type-A haemophilia.
XX Unidentified.
XX CN1361178-A.
XX 31-JUL-2002.
XX 29-DEC-2000; 2000CN-00137779.
XX 29-DEC-2000; 2000CN-00137779.
XX (SHAN-) SHANGHAI BIO-CHEM INST CHINESE ACAD SCI.
XX Qi Z, Wang Q, Chen C;
XX WPI; 2002-741852/81.
XX N-PSDB; ADE64593.
XX New recombinant blood coagulation factor VIII and its production process
XX PT and medicinal composition.
XX Claim 2; SEQ ID NO 4; 31pp; Chinese.
XX The invention relates to a novel recombinant blood coagulation factor
XX VIII, its production process and its medicinal composite for treating
XX type-A haemophilia. The invention further comprises a medicinal
XX composition containing the blood coagulation factor which promotes blood
XX coagulation to the blood plasma of type-A haemophilia patients. This
XX sequence represents a recombinant blood coagulation factor VIII protein
XX of the invention.

XX SQ Sequence 790 AA;
Query Match 100.0%; Score 871; DB 5; Length 790;
Best Local Similarity 100.0%; Pred. No. 3.9e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CDLNSCSMPLGWESKAISDAQITASSYFTNNFATWSPSKARLHLQGRSNARVQVNNPKE 60
Db 627 CDLNSCSMPLGWESKAISDAQITASSYFTNNFATWSPSKARLHLQGRSNARVQVNNPKE 686
QY 61 WLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 687 WLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 746
QY 121 SFTPVVNSLDPPLLRYLRIRHPSQSWHQIALRMEVLGCEAODLY 164
Db 747 SFTPVVNSLDPPLLRYLRIRHPSQSWHQIALRMEVLGCEAODLY 790

RESULT 5
AAP71139
ID AAP71139 standard; protein; 1014 AA.
XX

AAP71139;
25-MAR-2003 (revised)
14-MAY-1991 (first entry)
Factor VIII:c variant.
Factor VIII:c; variant; proteolysis; resistance;
pro coagulation activity.
Homo sapiens.
W08707144-A.
03-DEC-1987.
29-MAY-1987; 87WO-US001299.
29-MAY-1986; 86US-00869410.
18-NOV-1986; 86US-00932767.
09-DEC-1986; 86US-00939658.
(GEMY) GENETICS INST INC.
Kaufman RJ, Pittman D, Toole JJ;
WPI; 1987-348539/49.
New deletion and replacement variants of factor 8:c - resistant to
proteolysis but retaining pro-coagulant activity, and new DNA coding
sequences.
Disclosure; Page 1; 42pp; English.
The full-length human factor VIII:c cDNA has been set forth in detail in
W08501961. This sequences is an example of the formula A-X-B, wherein
A-Ala1-Arg372, B-Ser1690-Tyr2332 and X=0-1316 amino acids substantially
duplicative of sequences of amino acids within the sequence Arg372-
Ser1690 of the full-length sequence. Here X=0; producing a Arg371-Ser1690
fusion protein, one or more deletions or replacements at Arg 220, 226,
279, 282, 336, 359, 1719, 1721 and Lys 325, 338 and Tyr 346 will produce
variants which are more resistant to specific proteolytic cleavage
compared with natural factor VIII:c. Pro-coagulant activity and thrombin
activatability have been retained. See also AAP71726-29. (Updated on 25-
MAR-2003 to correct PA field.)
XX SQ Sequence 1014 AA;
Query Match 100.0%; Score 871; DB 1; Length 1014;
Best Local Similarity 100.0%; Pred. No. 5.5e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CDLNSCSMPLGWESKAISDAQITASSYFTNNFATWSPSKARLHLQGRSNARVQVNNPKE 60
Db 851 CDLNSCSMPLGWESKAISDAQITASSYFTNNFATWSPSKARLHLQGRSNARVQVNNPKE 910
QY 61 WLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 911 WLQVDFQKTMKVTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 970
QY 121 SFTPVVNSLDPPLLRYLRIRHPSQSWHQIALRMEVLGCEAODLY 164
Db 971 SFTPVVNSLDPPLLRYLRIRHPSQSWHQIALRMEVLGCEAODLY 1014
RESULT 6
AAW33227
ID AAW33227 standard; protein; 1383 AA.
XX AAW33227;
XX 30-APR-1998 (first entry)
XX

DE Procoagulant-active human factor VIII:C (FVIII) mutant protein.

XX Pro-coagulant active factor VIII; FVIII; haemophilia A;

KW recombinant secretion; pro-coagulant activity; resistance;

KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;

KW von Willebrand factor binding site; binding affinity;

KW FVIII replacement therapy.

XX Synthetic.

OS Homo sapiens.

XX Key

FT Region Location/Qualifiers

FT 1..346 /note= "factor VIIIA heavy chain"

FT Domain 1..329 /note= "A1 domain"

FT Domain 1..179 /note= "plastocyanin-like domain 1"

FT Disulfide-bond 153..179 /note= "probable"

FT Domain 187..329 /note= "plastocyanin-like domain 2"

FT Cleavage-site 372..373 /note= "by thrombin"

FT Domain 380..711 /note= "A2 domain"

FT Domain 380..554 /note= "A3 domain"

FT Disulfide-bond 528..554 /note= "probable"

FT Domain 564..711 /note= "plastocyanin-like domain 4"

FT Misc_feature 711..746 /note= "A spacer of the sequence SFQNSRHSSTKQKQFNATIPENDIEKTDWPF AHRTPMPKIQNVSSDILLMLL is inserted between domains A2 and A3"

FT Misc-difference 740 /label= "R740A"

FT Region /note= "wild type Arg replaced with Ala"

FT 741..1383 /note= "factor VIIIA light chain"

FT Domain 746..1073 /note= "A3 domain"

FT Domain 1073..1221 /note= "C1 domain"

FT Domain 1226..1378 /note= "C2 domain"

FT WO9740145-A1.

XX 30-OCT-1997.

XX 24-APR-1997; 97WO-US006563.

XX 24-APR-1996; 96US-0016117P.

PR 15-MAY-1996; 96US-0017785P.

XX (UNMI) UNIV MICHIGAN.

XX Kaufman RJ, Pipe SW, Amano K;

XX WPI; 1997-535830/49.

XX Modified human pro-coagulant active factor VIII - can be administered to haemophiliacs, i.e. factor VIII replacement therapy.

XX Claim 20; Page: 57pp; English.

XX The present sequence represents a novel pro-coagulant active factor VIII (FVIII) mutant protein, comprising a deletion of the B domain and von Willebrand factor binding site, a mutation at Arg740 and an addition of an amino acid sequence spacer between the A2 and A3 domains. Factor VIII, along with calcium and phospholipid, acts as a cofactor for factor IXa,

CC when it converts factor X to the activated form (factor XA). FVIII is the coagulation factor deficient in the X-chromosome-linked bleeding disorder haemophilia A. Several other mutant FVIII proteins have also been created (see AAW33222-29). The FVIII mutant F309S (AAW33225) is capable of recombinant secretion at higher levels than typically obtained with wild type FVIII and retains pro-coagulant activity. The FVIII mutant R336I (AAW33222) and R562K (AAW33223) are resistant to activated protein C (APC) cleavage. The present FVIII mutant can form a more stable configuration, and have an approximate 5-fold increase in specific activity compared to purified wild type FVIII, while increasing their binding affinity to von Willebrand factor improves their stability. The FVIII proteins can be administered to haemophiliacs, i.e. FVIII replacement therapy, while the nucleic acid molecule can be used for gene therapy. note: this sequence does not appear in the specification; it was created using sequences from the given references

XX Sequence 1383 AA;

SQ Query Match 100.0%; Score 871; DB 2; Length 1383;

Best Local Similarity 100.0%; Pred. No. 8.6e-90;

Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCSMPLEGMESKALSDAQITASSYFTNMFATWSPSKARLHLOGRSNARVQVNNPKE 60

Db 1220 CDLNSCSMPLEGMESKALSDAQITASSYFTNMFATWSPSKARLHLOGRSNARVQVNNPKE 1279

QY 61 WLQVDFQTKMKVTGTTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQD 120

Db 1280 WLQVDFQTKMKVTGTTQGVKSLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQD 1339

QY 121 SFTPEVNSLDPPLLTRYLRHPQSWHQAIALRMEVLGCEAQDLY 164

Db 1340 SFTPEVNSLDPPLLTRYLRHPQSWHQAIALRMEVLGCEAQDLY 1383

RESULT 7

AAW33228

ID AAW33228 standard; protein; 1383 AA.

XX AAW33228;

AC AAW33228;

DT 30-APR-1998 (first entry)

XX Procoagulant-active human factor VIII:C (FVIII) mutant protein.

DE Pro-coagulant active factor VIII; FVIII; haemophilia A;

KW recombinant secretion; pro-coagulant activity; resistance;

KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;

KW von Willebrand factor binding site; binding affinity;

KW FVIII replacement therapy.

XX Synthetic.

OS Homo sapiens.

XX Key

FT Region Location/Qualifiers

FT 1..346 /note= "factor VIIIA heavy chain"

FT Domain 1..329 /note= "A1 domain"

FT Domain 1..179 /note= "plastocyanin-like domain 1"

FT Disulfide-bond 153..179 /note= "probable"

FT Domain 187..329 /note= "plastocyanin-like domain 2"

FT Misc-difference 336 /label= "R336I"

FT /note= "wild type Arg replaced with Ile"

FT Cleavage-site 372..373 /note= "by thrombin"

FT Domain 380..711 /note= "A2 domain"

FT Domain 380..554 /note= "A3 domain"

FT /note= "plastocyanin-like domain 3"
 FT Disulfide-bond 528..554
 FT /note= "probable"
 FT Misc-difference 562
 FT /label= R562K
 FT /note= "wild type Arg replaced with Lys"
 FT 554..711
 FT Domain
 FT /note= "plastocyanin-like domain 4"
 FT 711..746
 FT Misc_feature
 FT /note= "a spacer of the sequence
 FT SFQSRHPSTKQFNATIPENDIKDPWF AHRTPMPKIQNVSSDILLMLL
 FT is inserted between domains A2 and A3"
 FT Misc-difference 740
 FT /label= R740A
 FT /note= "wild type Arg replaced with Ala"
 FT 741..1383
 FT Region
 FT /note= "factor VIIIA light chain"
 FT 746..1073
 FT Domain
 FT /note= "A3 domain"
 FT 1073..1221
 FT Domain
 FT /note= "C1 domain"
 FT 1226..1378
 FT Domain
 FT /note= "C2 domain"
 FT
 FT WO9740145-A1.
 PN
 XX
 XX 30-OCT-1997.
 PD
 XX
 XX 24-APR-1997; 97WO-US006563.
 XX
 XX 24-APR-1996; 96US-0016117P.
 PR
 PR 15-MAY-1996; 96US-0017785P.
 XX
 XX (UNMI) UNIV MICHIGAN.
 PA
 XX Kaufman RJ, Pipe SW, Amano K;
 PI
 PI WPI; 1997-535830/49.
 DR
 XX
 XX Modified human pro-coagulant active factor VIII - can be administered to
 PT haemophiliacs, i.e. factor VIII replacement therapy.
 PT
 XX Claim 18; Page; 57pp; English.
 PS
 XX
 XX The present sequence represents a novel pro-coagulant active factor VIII
 CC (FVIII) mutant protein, comprising a deletion of the B domain and von
 CC Willebrand factor binding site, mutations R336I, R582K and R740A and an
 CC addition of an amino acid sequence spacer between the A2 and A3 domains.
 CC Factor VII, along with calcium and phospholipid, acts as a cofactor for
 CC factor IXA, when it converts factor X to the activated form (factor XA).
 CC FVIII is the coagulation factor deficient in the X-chromosome-linked
 CC bleeding disorder haemophilia A. Several other mutant FVIII proteins have
 CC also been created (see AAW33222-29). The FVIII mutant F309S (AAW33225) is
 CC capable of recombinant secretion at higher levels than typically obtained
 CC with wild type FVIII and retains pro-coagulant activity. The FVIII mutant
 CC R336I (AAW33222) and R562K (AAW33223) are resistant to activated protein
 CC C (APC) cleavage. The FVIII mutant comprising a deletion of the B domain
 CC and von Willebrand factor binding site, a mutation at Arg740 and an
 CC addition of an amino acid sequence spacer between the A2 and A3 domains
 CC can form a more stable configuration, and have an approximate 5-fold
 CC increase in specific activity compared to purified wild type FVIII, while
 CC increasing their binding affinity to von Willebrand factor improves their
 CC stability. The FVIII proteins can be administered to haemophiliacs, i.e.
 CC FVIII replacement therapy, while the nucleic acid molecule can be used
 CC for gene therapy. note: this sequence does not appear in the
 CC specification; it was created using sequences from the given references
 XX
 XX Sequence 1383 AA;
 SQ

Query Match 100.0%; Score 871; DB 2; Length 1383;
 Best Local Similarity 100.0%; Pred. No. 8,6e-30;
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CDLNSCSNPLGMSKALSDAQITASSYFTNMFATWSKARLHLOGRSNARPOVNNPKE 60
 Db 1220 CDLNSCSNPLGMSKALSDAQITASSYFTNMFATWSKARLHLOGRSNARPOVNNPKE 1279
 Oy 61 WLQVDFOKTMKVTGTTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFOGNO 120
 Db 1280 WLQVDFOKTMKVTGTTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFOGNO 1339
 Oy 121 SFTPVNSLPPPLTLRLRHPSQSWHQAIRMEVLGCEAODLY 164
 Db 1340 SFTPVNSLPPPLTLRLRHPSQSWHQAIRMEVLGCEAODLY 1383
 RESULT 8
 AAW33229
 ID AAW33229 standard; protein; 1383 AA.
 XX
 AC AAW33229;
 XX
 DT 30-APR-1998 (first entry)
 XX
 DE Procoagulant-active human factor VIII:C (FVIII) mutant protein.
 XX
 KW Pro-coagulant active factor VIII; FVIII; haemophilia A;
 KW recombinant secretion; pro-coagulant activity; resistance;
 KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
 KW von Willebrand factor binding site; binding affinity;
 KW FVIII replacement therapy.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key
 FT Region
 FT /note= "factor VIIIA heavy chain"
 FT Domain
 FT /note= "A1 domain"
 FT Domain
 FT /note= "plastocyanin-like domain 1"
 FT Disulfide-bond 153..179
 FT /note= "probable"
 FT Domain
 FT /note= "plastocyanin-like domain 2"
 FT Misc-difference 309
 FT /label= F309S
 FT /note= "wild type Phe replaced with Ser"
 FT Cleavage-site 372..373
 FT /note= "by thrombin"
 FT Domain
 FT /note= "A2 domain"
 FT Domain
 FT /note= "plastocyanin-like domain 3"
 FT Disulfide-bond 380..554
 FT /note= "probable"
 FT Domain
 FT /note= "plastocyanin-like domain 4"
 FT Misc_feature 711..746
 FT /note= "a spacer of the sequence
 FT SFQSRHPSTKQFNATIPENDIKDPWF AHRTPMPKIQNVSSDILLMLL
 FT is inserted between domains A2 and A3"
 FT Misc-difference 740
 FT /label= R740A
 FT /note= "wild type Arg replaced with Ala"
 FT 741..1383
 FT Region
 FT /note= "factor VIIIA light chain"
 FT Domain
 FT /note= "A3 domain"
 FT 1073..1221
 FT Domain
 FT /note= "C1 domain"
 FT 1226..1378
 FT Domain
 FT /note= "C2 domain"
 XX
 PN WO9740145-A1.

XX PD 30-OCT-1997.
 XX PF 24-APR-1997; 97WO-US006563.
 XX PR 24-APR-1996; 96US-0016117P.
 XX PR 15-MAY-1996; 96US-0017785P.
 XX PA (UNMI) UNIV MICHIGAN.
 XX KA Kaufman RJ, Pipe SW, Amano K;
 XX WPI; 1997-535830/49.
 XX PT Modified human pro-coagulant active factor VIII - can be administered to
 XX PT haemophiliacs, i.e. factor VIII replacement therapy.
 XX PS Claim 19; Page; 57pp; English.
 XX CC The present sequence represents a novel pro-coagulant active factor VIII
 XX CC (FVIII) mutant protein, comprising a deletion of the B domain and von
 XX CC Willebrand factor binding site, mutations F309S, R740A and addition of an
 XX CC amino acid sequence spacer between the A2 and A3 domains. Factor VIII,
 XX CC along with calcium and phospholipid, acts as a cofactor for factor IXA,
 XX CC when it converts factor X to the activated form (factor Xa). FVIII is the
 XX CC coagulation factor deficient in the X-chromosome-linked bleeding disorder
 XX CC haemophilia A. Several other mutant FVIII proteins have also been created
 XX CC (see AAW33222-29). The FVIII mutant F309S (AAW33225) is capable of
 XX CC recombinant secretion at higher levels than typically obtained with wild
 XX CC type FVIII and retains pro-coagulant activity. The FVIII mutant R336I
 XX CC (AAW33222) and R562K (AAW33223) are resistant to activated protein C
 XX CC (APC) cleavage. The present FVIII mutant can form a more stable
 XX CC configuration, and have an approximate 5-fold increase in specific
 XX CC activity compared to purified wild type FVIII, while increasing their
 XX CC binding affinity to von Willebrand factor improves their stability. The
 XX CC FVIII proteins can be administered to haemophiliacs, i.e. FVIII
 XX CC replacement therapy, while the nucleic acid molecule can be used for Gene
 XX CC therapy. note: this sequence does not appear in the specification; it was
 XX CC created using sequences from the given references
 XX SQ Sequence 1383 AA;
 Query Match 100.0%; Score 871; DB 2; Length 1383;
 Best Local Similarity 100.0%; Pred. No. 8.6e-90;
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CDLNSCSMPGLMESKAISDAQITASSYFTNMFTWSPSKARLHLOGRSNAMRPQVNNPKE 60
 DB 1220 CDLNSCSMPGLMESKAISDAQITASSYFTNMFTWSPSKARLHLOGRSNAMRPQVNNPKE 1279
 QY 61 WLQVDFQKTKMKTGVTGTTQGVKSLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
 DB 1280 WLQVDFQKTKMKTGVTGTTQGVKSLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQD 1339
 QY 121 SFTPVNSLDPPLLTRYLRHPSQSWHQIALRMEVLGCEAQDLY 164
 DB 1340 SFTPVNSLDPPLLTRYLRHPSQSWHQIALRMEVLGCEAQDLY 1383
 RESULT 9
 AAP80268
 ID AAP80268 standard; protein; 1424 AA.
 XX AC AAP80268;
 XX DT 25-MAR-2003 (revised)
 XX DT 10-OCT-1990 (first entry)
 XX DE Modified factor VIII:C sequence with the R740-D1658 deletion.
 XX KW Modified factor VIII:C; maturation polypeptide; haemophilia;
 XX KW blood coagulation; RD deletion; procoagulant.
 XX XX

OS Homo sapiens.
 XX WO8800831-A.
 XX PD 11-FEB-1988.
 XX PF 31-JUL-1987; 87WO-US001814.
 XX PR 01-AUG-1986; 86US-00893375.
 XX PA (BIOJ) BIOGEN NV.
 XX PA (PASE/) PASEK M P.
 XX PI Pasek MP;
 XX WPI; 1988-049866/07.
 XX DR N-PSDB; AAN80447.
 XX PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA
 XX PT encoding maturation polypeptide, useful for high yield transformation.
 XX PS Claim 3; Page 60-61-62-63; 97pp; English.
 XX CC The RD deletion removes the DNA from Ser 741 to Ser 1657. A major part of
 XX CC the sequence encoding the maturation polypeptide of factor VIII:C is
 XX CC deleted, i.e. Gln 744 - Asp 1563. The full length Factor VIII:C cDNA has
 XX CC two changes with respect to the published sequence (EPO application
 XX CC 160457): CTG to CTA at Leu 242 and TTC to CTC change at amino acid
 XX CC residue 1880 (Phe to Leu). The product is produced in approx. 20 times
 XX CC higher yields than previous recombinant produced factor VIII:C and are
 XX CC more easily purified. The peptide is used for treating haemophilia A,
 XX CC both acute and prolonged bleeding. See also AAN80444 and AAN80446.
 XX CC (Updated on 25-MAR-2003 to correct PA field.)
 XX SQ Sequence 1424 AA;
 Query Match 100.0%; Score 871; DB 1; Length 1424;
 Best Local Similarity 100.0%; Pred. No. 9e-90;
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CDLNSCSMPGLMESKAISDAQITASSYFTNMFTWSPSKARLHLOGRSNAMRPQVNNPKE 60
 DB 1261 CDLNSCSMPGLMESKAISDAQITASSYFTNMFTWSPSKARLHLOGRSNAMRPQVNNPKE 1320
 QY 61 WLQVDFQKTKMKTGVTGTTQGVKSLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
 DB 1321 WLQVDFQKTKMKTGVTGTTQGVKSLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQD 1380
 QY 121 SFTPVNSLDPPLLTRYLRHPSQSWHQIALRMEVLGCEAQDLY 164
 DB 1381 SFTPVNSLDPPLLTRYLRHPSQSWHQIALRMEVLGCEAQDLY 1424
 RESULT 10
 AAP91169
 ID AAP91169 standard; protein; 1424 AA.
 XX AC AAP91169;
 XX DT 25-MAR-2003 (revised)
 XX DT 26-JUN-1990 (first entry)
 XX DE Sequence of 740 Arg-1649 Glu human Factor VIII:C.
 XX KW Human Factor VIII:C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII:C;
 XX KW haemophilia A.
 XX OS Homo sapiens.
 XX XX EF306968-A.
 XX PD 15-MAR-1989.
 XX XX

PF 09-SEP-1988; 88EP-00114769.
XX 10-SEP-1987; 87JP-00225147.
PR 08-APR-1988; 88JP-00085454.
XX
PA (KAGA) CHEMO SERO THERAPEUTIC RES INST.
PA (TEIU) TEIJIN LTD.
XX
PI Sugiyama T, Masuda K, Tajima Y, Yonemura H;
XX WPI; 1989-078467/11.
DR N-PSDB; AAN90654.
XX
XX Prodn. of recombinant human Factor-VIII-C - using animal cells
PT transformed with a vector contg. the gene for Factor VIII:C and a
PT promoter.
XX
PS Disclosure; Fig 1; 32pp; English.
XX
CC Arg-740 of the carboxyl terminus of the H chain is directly bonded by a
CC peptide bond to Glu-1649 of the amino terminus of L chain. A prefd.
CC expression vector used to transform animal cell so that they produce
CC human Factor VIII:Cis plasmid Ad.RE.neo. The expression vector has at
CC least one promoter upstream of AAN90654. The transformants can constantly
CC and continuously produce human Factor VIII:C in high yield on a
CC commercial scale. The human Factor VIII:C so produced is considered to
CC corresp. to the smallest species of active and intact Factor VIII:C
CC molecules in the human blood plasma. It is useful for treating
CC haemophilia A patients. (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1424 AA;

Query Match 100.0%; Score 871; DB 1; Length 1424;
Best Local Similarity 100.0%; Pred. No. 9e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCSMP:GMESKAISDAQITASSYFTNMFTWSPSKARLHLQGRSNARWPQVNNPKE 60
DB 1261 CDLNSCSMP:GMESKAISDAQITASSYFTNMFTWSPSKARLHLQGRSNARWPQVNNPKE 1320

QY 61 WLQVDFQKTMKVTGVTQTGGVKSLLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
DB 1321 WLQVDFQKTMKVTGVTQTGGVKSLLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQD 1380

QY 121 SFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY 164
DB 1381 SFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY 1424

RESULT 11
AAB48842
ID AAB48842 standard; protein; 1424 AA.
XX
AC AAB48842;
XX
DT 13-MAR-2001 (first entry)
XX
DE Mutant mature human factor VIII, SEQ ID NO:5.
XX
XX Factor VIII; human; B domain; LRP-mediated plasma clearance; half-life;
KW receptor-dependent clearance; receptor-independent clearance; half-life;
KW haemophilia; mutant; mutein.
XX
OS Homo sapiens.
XX
XX WO200071714-A2.
XX
XX 30-NOV-2000.
XX
XX 24-MAY-2000; 2000WO-US014111.
XX
XX 24-MAY-1999; 99US-0135847P.
PR

XX (AMNA-) AMERICAN NAT RED CROSS.
PA Saenko EL, Strickland DK;
XX WPI; 2001-025163/03.
XX
FT Factor VIII mutants having increased half-life useful for treating
FT hemophilia, comprise one or more amino acid substitutions in the A2
FT and/or C2 domain of factor VIII.
XX
PS Claim 9; Fig 2A-B; 121pp; English.
XX
XX The invention relates to human factor VIII mutants comprising an amino
CC acid substitution at one or more positions in the A2 domain and/or an
CC amino acid substitution at one or more positions in the C2 domain. The
CC invention also encompasses a factor VIII mutant which lacks a B domain
CC (AAB48842). The factor VIII mutants have an increased half-life in the
CC bloodstream. The A2 domain mutants exhibit reduced LRP-dependent
CC (receptor-dependent) clearance of factor VIII, while C2 domain mutants
CC have reduced receptor-independent clearance. The invention also relates
CC to a method of using RAP (receptor associated protein), a protein which
CC inhibits LRP (low density lipoprotein related protein)-mediated ligand
CC internalisation, to increase the half-life of factor VIII. The mutant
CC factor VIII proteins, and nucleotides encoding them, are useful for
CC treating haemophilia. RAP, LRP-binding RAP mutants or fragments, and
CC nucleic acids encoding them may also be used in the treatment of
CC haemophilia, in combination with a mutant factor VIII protein or DNA of
CC the invention. The invention provides means of increasing the half-life
CC of factor VIII by reducing its clearance from plasma. The present
CC sequence represents a mutant mature human factor VIII which lacks a B
CC domain
XX
SQ Sequence 1424 AA;

Query Match 100.0%; Score 871; DB 4; Length 1424;
Best Local Similarity 100.0%; Pred. No. 9e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCSMP:GMESKAISDAQITASSYFTNMFTWSPSKARLHLQGRSNARWPQVNNPKE 60
DB 1261 CDLNSCSMP:GMESKAISDAQITASSYFTNMFTWSPSKARLHLQGRSNARWPQVNNPKE 1320

QY 61 WLQVDFQKTMKVTGVTQTGGVKSLLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
DB 1321 WLQVDFQKTMKVTGVTQTGGVKSLLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQD 1380

QY 121 SFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY 164
DB 1381 SFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY 1424

RESULT 12
AAO18622
ID AAO18622 standard; protein; 1424 AA.
XX
AC AAO18622;
XX
DT 24-OCT-2002 (first entry)
XX
DE Human mature B-domainless factor VIII.
XX
XX Human; factor VIII; fVIII; half-life; mutant; haemophilia;
KW heparan sulfate proteoglycan-mediated clearance; RAP;
KW receptor-associated protein; haemostatic; gene therapy;
KW alpha2 macroglobulin receptor-associated protein.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 373..740
FT Domain /label= A2_domain
FT Region 484..509
FT

```
FT XX /label= LRP_binding_region
PN XX WO200260951-A2.
PD XX
PP XX 08-AUG-2002.
XX XX
XX XX 11-JAN-2002; 2002WO-US000583.
XX XX
XX XX 12-JAN-2001; 2001US-0250904P.
XX XX (AMNA-) AMERICAN NAT RED CROSS.
XX XX
XX XX Saenko EL, Sarafanov AG;
XX XX WPI; 2002-608501/65.
XX XX
XX XX New mutant factor VIII with reduced sulfate proteoglycan (HSPG)-dependent
XX XX or receptor-independent clearance and procoagulant activity for treating
XX XX hemophilia.
XX XX
XX XX Claim 9; Fig 12; 161pp; English.
XX XX
XX XX The present invention relates to a mutant factor VIII protein with
XX XX reduced sulfate proteoglycan (HSPG)-dependent or receptor-independent
XX XX clearance and procoagulant activity, which has a nonconservative amino
XX XX acid substitution at one or more positions in the A2 domain consisting of
XX XX Lys(380, 512, 556, 570 or 659) or Arg(490, 527, 562 or 571) or in the C2
XX XX domain relative to the wild-type. The mutant factor VIII or the
XX XX polynucleotide encoding it and a receptor-associated protein (alpha2
XX XX macroglobulin receptor-associated protein or RAP) are useful for treating
XX XX haemophilia. The mutated protein has a longer half-life. The present
XX XX sequence is the human mature B-domainless factor VIII protein
XX XX
XX XX Sequence 1424 AA;
XX XX
XX XX Query Match 100.0%; Score 871; DB 5; Length 1424;
XX XX Best Local Similarity 100.0%; Pred. No. 9e-90;
XX XX Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX XX 1 CDLNSCSMPLGMSKALSDAQITASSYFTNMFTWSPSKARLHLQGRSNARVPQVNNPKE 60
XX XX |||||||
XX XX 1261 CDLNSCSMPLGMSKALSDAQITASSYFTNMFTWSPSKARLHLQGRSNARVPQVNNPKE 1320
XX XX |||||||
XX XX 61 WLQVDFQKTKMKTGVTGQVKSLLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
XX XX |||||||
XX XX 1321 WLQVDFQKTKMKTGVTGQVKSLLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQD 1380
XX XX |||||||
XX XX 121 SFTPVNSLDPPLLTLYRLRHPSQSWVHQIALRMEVLGCEAQDLY 164
XX XX |||||||
XX XX 1381 SFTPVNSLDPPLLTLYRLRHPSQSWVHQIALRMEVLGCEAQDLY 1424
XX XX |||||||
XX XX
XX XX RESULT 13
XX XX AAP80267
XX XX ID AAP80267 standard; protein; 1425 AA.
XX XX
XX XX AAP80267;
XX XX
XX XX 25-MAR-2003 (revised)
XX XX DT 10-OCT-1990 (first entry)
XX XX
XX XX Modified factor VIII:C sequence with the R740-E1649 deletion.
XX XX
XX XX Modified factor VIII:C; maturation polypeptide; haemophilia;
XX XX blood coagulation; RE deletion.
XX XX
XX XX Homo sapiens.
XX XX
XX XX WC8800831-A.
XX XX
XX XX 11-FEB-1988.
XX XX
XX XX 31-JUL-1987; 87WO-US001814.
XX XX
```

```
XX XX
XX XX 01-AUG-1986; 86US-00893375.
XX XX (BIOJ ) BIOGEN NV.
XX XX (PASE/) PASEK M P.
XX XX
XX XX Pasek MP;
XX XX
XX XX WPI; 1988-C49866/07.
XX XX N-PSDB; AAN80446.
XX XX
XX XX New DNA sequences encoding modified factor VIII:C - with deletion of DNA
XX XX encoding maturation polypeptide, useful for high yield transformation.
XX XX
XX XX Claim 3; Page 57-58-59-60; 97pp; English.
XX XX
XX XX The entire sequence encoding the maturation polypeptide of factor VIII:C
XX XX is deleted, i.e. Arg 740-Glu 1649. The full length factor VIII:C cDNA has
XX XX two changes with respect to the published sequence (BPO application
XX XX 160457): CTG to CTA at Leu 242 and TTC to CTC change at amino acid
XX XX residue 1880 (Phe to Leu). The product is produced in approx. 20 times
XX XX higher yields than previous recombinant produced factor VIII:C and are
XX XX more easily purified. The peptide is used for treating haemophilia A,
XX XX both acute and prolonged bleeding. See also AAN80444 and AAN80447.
XX XX (Updated on 25-MAR-2003 to correct PA field.)
XX XX
XX XX Sequence 1425 AA;
XX XX
XX XX Query Match 100.0%; Score 871; DB 1; Length 1425;
XX XX Best Local Similarity 100.0%; Pred. No. 9e-90;
XX XX Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX XX 1 CDLNSCSMPLGMSKALSDAQITASSYFTNMFTWSPSKARLHLQGRSNARVPQVNNPKE 60
XX XX |||||||
XX XX 1262 CDLNSCSMPLGMSKALSDAQITASSYFTNMFTWSPSKARLHLQGRSNARVPQVNNPKE 1321
XX XX |||||||
XX XX 61 WLQVDFQKTKMKTGVTGQVKSLLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
XX XX |||||||
XX XX 1322 WLQVDFQKTKMKTGVTGQVKSLLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGNQD 1381
XX XX |||||||
XX XX 121 SFTPVNSLDPPLLTLYRLRHPSQSWVHQIALRMEVLGCEAQDLY 164
XX XX |||||||
XX XX 1382 SFTPVNSLDPPLLTLYRLRHPSQSWVHQIALRMEVLGCEAQDLY 1425
XX XX |||||||
XX XX
XX XX RESULT 14
XX XX AAB01262
XX XX ID AAB01262 standard; protein; 1438 AA.
XX XX
XX XX AAB01262;
XX XX
XX XX 25-SEP-2000 (first entry)
XX XX
XX XX B-domain deleted factor VIII sequence.
XX XX
XX XX Factor VIII; procoagulant; adenovirus; adeno-associated strain;
XX XX Gene therapy; human Burkitt's lymphoma; HKB; therapy;
XX XX therapeutic protein; vector; Epstein-Barr virus; human.
XX XX
XX XX Homo sapiens.
XX XX
XX XX WO200034505-A1.
XX XX
XX XX 15-JUN-2000.
XX XX
XX XX 08-DEC-1999; 99WO-US029169.
XX XX
XX XX 10-DEC-1998; 98US-00209916.
XX XX (FARB ) BAYER CORP.
XX XX
XX XX Cho M, Chan SY, Kelsey W, Yee H;
XX XX
```

DR WPI; 2000-431311/37.
XX
PT Producing cells expressing a protein having factor VIII procoagulant
PT activity especially, human factor VIII in an industrial scale, involves
PT expressing a vector comprising a sequence coding for factor VIII in human
PT cells.
XX
XX Claim 7; Fig 1; 27pp; English.
PS
XX Producing cells expressing a protein having factor VIII procoagulant
CC activity, comprises contacting the cells with a vector comprising a
CC selectable marker and a sequence coding for the protein having factor
CC VIII procoagulant activity operably linked to a promoter. The cells are
CC then selected and individual clones expressing high levels of the protein
CC are isolated from the selected cells. The cells produced by the method
CC are not only useful for producing protein having factor VIII procoagulant
CC activity but also for producing adenovirus and adeno-associated virus
CC strains for gene therapy. The advantage of having cells producing protein
CC with factor VIII procoagulant activity is that factor VIII protein can be
CC produced on an industrial scale in the range of 2-4 pg/cell/day. Human
CC Burkitt's lymphoma (HKB) cells provide a protein-free production system
CC to produce not only B-domain deleted factor VIII but also other
CC therapeutic proteins. The vector used in the method preferably comprises
CC B-domain deleted factor VIII (BDD-FVIII), a transcriptional unit for BDD-
CC FVIII and a selectable marker, dihydrofolate reductase (dhfr). In
CC addition, a terminal repeat sequence from Epstein-Barr virus is inserted
CC into the vector to increase integration efficiency
XX
XX Sequence 1438 AA;
Query Match 100.0%; Score 871; DB 3; Length 1438;
Best Local Similarity 100.0%; Pred. No. 9.1e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CDLNSCSMPGLGMSKAIISDAQITASSYFTNMFATWSPSKARLHLQGRSNARWPQVNNPKE 60
DB 1275 CDLNSCSMPGLGMSKAIISDAQITASSYFTNMFATWSPSKARLHLQGRSNARWPQVNNPKE 1334
QY 61 WLQVDFQKTKVGTGTTQGVKSLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
DB 1335 WLQVDFQKTKVGTGTTQGVKSLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQD 1394
QY 121 SFTPVNSLDPEPLLTRYLRIHPQSWHQAIALRMEVLGCEAODLY 164
DB 1395 SFTPVNSLDPEPLLTRYLRIHPQSWHQAIALRMEVLGCEAODLY 1438
RESULT 15
AAR12971
XX AAR12971 standard; protein; 1440 AA.
XX
XX AAR12971;
XX
XX 25-MAR-2003 (revised)
DT 09-JAN-2003 (revised)
DT 02-OCT-1991 (first entry)
XX
XX Factor VIII:SQ.
XX
XX Factor VIII; B domain; haemophilia.
XX
XX Unidentified.
XX
XX WO9109122-A.
XX
XX 27-JUN-1991.
XX
XX 15-DEC-1989; 89SE-00004239.
XX
XX 15-DEC-1989; 89SE-00004239.
XX
XX (KABI) KABIVITRUM AB.

PI Almstedt AB, Hellstrom EM, Larsson K, Lind P, Sandberg HI;
PI Spira J, Sydowback WM;
XX WPI; 1991-208148/28.
DR Recombinant human factor VIII deriv. deoxyribonucleic acid - encoding
XX protein comprising two chains linked by segment of B domain.
PT
XX Disclosure; Fig 1; 35pp; English.
PS
XX The protein is a fusion between Phe 742 and Ser 1637 of the factor VIII
CC protein (factor VIII:SQ). In order to produce a factor VIII deletion
CC derivative that can be produced in vivo and/or in vitro, to a two chain
CC protein consisting of polypeptide chains of 90 kD and 80 kD, the amino
CC acid sequences surrounding Arg 740 and Arg 1648 have to be conserved in
CC order to preserve the structural requirements for correct cleavage. In
CC this example, amino acids 743 to 1636 of the full-length factor VIII
CC polypeptide are deleted. A new polypeptide chain is obd. where there are
CC 14 amino acids linking Arg 740 and Arg 1648. Of these 14 amino acids, the
CC sequence of the five N-terminal ones directly corresponds to the five
CC amino acids following Arg 740 in full-length factor VIII. Also, the
CC sequence of the 12 C-terminal amino acids of the above 14 amino acids
CC fragment directly corresponds to the 12 amino acids preceding Glu 1699 in
CC full-length factor VIII, thus creating a 3 amino acid overlap between
CC the N- and C-terminal regions of the B-domain. The factor VIII deriv. is
CC useful for treating haemophilia or haemophilia A. It has the biological
CC characteristics of plasma derived factor VIII. In order to index this
CC example, the factor VIII:Qb amino acid sequence was retrieved from
CC WO800831 (AAP80265). The amino acid numbering in the above comments is
CC reproduced from the fig. description in the specification. Note that Arg
CC 740 is Arg 742 in the indexed sequence, etc. Also, Asn 745 (N-terminal
CC link overlap) is Asp 747 in AAP80265, but indexed as Asn to reproduce the
CC fusion fragment as shown in fig.1, and Asn 745 (C-terminal link overlap).
CC (Updated on 09-JAN-2003 to add missing OS field.) (Updated on 25-MAR-2003
CC to correct PI field.)
XX
XX Sequence 1440 AA;
Query Match: 100.0%; Score 871; DB 2; Length 1440;
Best Local Similarity 100.0%; Pred. No. 9.2e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CDLNSCSMPGLGMSKAIISDAQITASSYFTNMFATWSPSKARLHLQGRSNARWPQVNNPKE 60
DB 1277 CDLNSCSMPGLGMSKAIISDAQITASSYFTNMFATWSPSKARLHLQGRSNARWPQVNNPKE 1336
QY 61 WLQVDFQKTKVGTGTTQGVKSLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
DB 1337 WLQVDFQKTKVGTGTTQGVKSLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQD 1396
QY 121 SFTPVNSLDPEPLLTRYLRIHPQSWHQAIALRMEVLGCEAODLY 164
DB 1397 SFTPVNSLDPEPLLTRYLRIHPQSWHQAIALRMEVLGCEAODLY 1440
Search completed: September 29, 2004, 16:38:58
Job time : 58 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 16:40:09 ; Search time 131 Seconds
(without alignments)
402.863 Million cell updates/sec

Title: US-10-049-3999a-1_COPY_2169_2332

Perfect score: 871
Sequence: 1 CDLNSCSMPLGMSKAISDA.....VWQIALRMEVLGCEAQDLY 164

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	871	100.0	1438	13 US-10-006-091-1
2	871	100.0	1438	13 US-10-047-257-1
3	871	100.0	1438	14 US-10-225-900-1
4	871	100.0	1459	16 US-10-239-498A-13
5	871	100.0	1471	12 US-10-681-970-2
6	871	100.0	1471	13 US-10-095-718-2
7	871	100.0	2332	9 US-09-957-641-2
8	871	100.0	2332	14 US-10-187-319-2
9	871	100.0	2332	14 US-10-131-510A-2
10	871	100.0	2332	15 US-10-445-235-2
11	871	100.0	2332	15 US-10-360-101-229
12	871	100.0	2332	16 US-10-239-498A-2
13	871	100.0	2332	16 US-10-466-998A-1
14	871	100.0	2351	12 US-10-411-037-30
15	871	100.0	2351	12 US-10-411-026-30

16	871	100.0	2351	14 US-10-132-829-4
17	871	100.0	2351	14 US-10-172-712-27
18	871	100.0	2351	14 US-10-133-907-4
19	871	100.0	2351	16 US-10-410-962-30
20	871	100.0	2351	16 US-10-411-049-30
21	871	100.0	2351	16 US-10-410-930-30
22	871	100.0	2351	16 US-10-410-997-30
23	871	100.0	2351	16 US-10-411-012-30
24	871	100.0	2351	16 US-10-287-994-30
25	871	100.0	2351	16 US-10-410-913-30
26	867	99.5	218	16 US-10-038-252-5
27	865	99.3	1459	16 US-10-239-498A-4
28	865	99.3	1459	16 US-10-239-498A-15
29	846	97.1	160	15 US-10-298-796-9
30	752	86.3	2319	14 US-10-187-319-6
31	752	86.3	2319	14 US-10-131-510A-6
32	736	84.5	1431	12 US-10-681-970-4
33	736	84.5	1431	13 US-10-095-718-4
34	720	82.7	1443	14 US-10-187-319-39
35	720	82.7	1443	14 US-10-131-510A-39
36	720	82.7	2133	14 US-10-187-319-37
37	720	82.7	2133	14 US-10-131-510A-37
38	683	78.4	130	15 US-10-298-796-1
39	383	44.0	80	16 US-10-433-273-7
40	380	43.6	218	16 US-10-038-252-4
41	380	43.6	2224	16 US-10-741-601-542
42	376	43.2	2196	15 US-10-360-101-259
43	376	43.2	2224	14 US-10-115-563-14
44	376	43.2	2224	14 US-10-172-712-31
45	366.5	42.1	480	14 US-10-177-293-122

ALIGNMENTS

RESULT 1

US-10-006-091-1

; Sequence 1, Application US/10006091

; Publication No. US20020102730A1

; GENERAL INFORMATION:

; APPLICANT: Cho, Myung-Sam

; APPLICANT: Chan, Sham-Yuen

; APPLICANT: Kelsey, William

; APPLICANT: Yee, Helena

; TITLE OF INVENTION: Expression System for Factor VIII

; FILE REFERENCE: MSB-7255.1

; CURRENT APPLICATION NUMBER: US/10/006,091

; CURRENT FILING DATE: 2001-12-06

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 1438

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Derived from

; OTHER INFORMATION: human factor VIII sequence

US-10-006-091-1

Query Match	100.0%	Score 871;	DB 13;	Length 1438;
Best Local Similarity	100.0%;	Pred. No. 2.4e-90;		
Matches 164;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CDLNSCSMPLGMSKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNARWPQVNNPK 60		
Db	1275	CDLNSCSMPLGMSKAISDAQITASSYFTNMFATWSPSKARLHLOGRSNARWPQVNNPK 1334		
QY	61	WLQVDFQKTKYTGVTGQVKSLLTSVMYKEFLISSQDGHQHTLFFQNGKVKVFQGNQD 120		
Db	1335	WLQVDFQKTKYTGVTGQVKSLLTSVMYKEFLISSQDGHQHTLFFQNGKVKVFQGNQD 1394		
QY	121	SFTPVVNSLDPPLLRYLRIHFQSWVHQIALRMEVLGCEAQDLY 164		

Db 1395 SFTPVNSLDPLLLTRYLRHQPQSWHQAIALRMEVLGCEAODLY 1438

RESULT 2
US-10-047-257-1
; Sequence 1, Application US/10047257
; Publication No. US2002011512A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255.2
; CURRENT APPLICATION NUMBER: US/10/047,257
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-047-257-1

Query Match 100.0%; Score 871; DB 13; Length 1438;
Best Local Similarity 100.0%; Pred. No. 2.4e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDLNSCSMPLGMSKAIISDAQITASSYFTNMFTWSPSKARLHLOGRSNAWRPQVNNPKE 60
Db 1275 CDLNSCSMPLGMSKAIISDAQITASSYFTNMFTWSPSKARLHLOGRSNAWRPQVNNPKE 1334

Qy 61 WLQVDFQKTMKVGTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 1335 WLQVDFQKTMKVGTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 1394

Qy 121 SFTPVNSLDPLLLTRYLRHQPQSWHQAIALRMEVLGCEAODLY 164
Db 1395 SFTPVNSLDPLLLTRYLRHQPQSWHQAIALRMEVLGCEAODLY 1438

RESULT 4
US-10-239-498A-13
; Sequence 13, Application US/10239498A
; Publication No. US2004002333A1
; GENERAL INFORMATION:
; APPLICANT: Hauser, Charlotte
; APPLICANT: Horster, Andrea
; APPLICANT: Schroder, Carola
; APPLICANT: Lehner, Michael
; TITLE OF INVENTION: Production of Recombinant Blood Clotting Factors in
; TITLE OF INVENTION: Human Cell Lines
; FILE REFERENCE: 80977.0001
; CURRENT APPLICATION NUMBER: US/10/239,498A
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: PCT/EP01/03220
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1459
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pTGF8-2hyg-s
US-10-239-498A-13

Query Match 100.0%; Score 871; DB 16; Length 1459;
Best Local Similarity 100.0%; Pred. No. 2.4e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDLNSCSMPLGMSKAIISDAQITASSYFTNMFTWSPSKARLHLOGRSNAWRPQVNNPKE 60
Db 1296 CDLNSCSMPLGMSKAIISDAQITASSYFTNMFTWSPSKARLHLOGRSNAWRPQVNNPKE 1355

Qy 61 WLQVDFQKTMKVGTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 1356 WLQVDFQKTMKVGTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 1415

Qy 121 SFTPVNSLDPLLLTRYLRHQPQSWHQAIALRMEVLGCEAODLY 164
Db 1416 SFTPVNSLDPLLLTRYLRHQPQSWHQAIALRMEVLGCEAODLY 1459

RESULT 5
US-10-681-970-2
; Sequence 2, Application US/10681970
; Publication No. US20040062752A1
; GENERAL INFORMATION:
; APPLICANT: Walsh, Christopher
; APPLICANT: Chao, Hengjun
; APPLICANT: Burstein, Haim
; APPLICANT: Lynch, Carmel
; APPLICANT: Stepan, Tony
; APPLICANT: Mursch, Keith
; TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
; FILE REFERENCE: Methods of Using the Same
; FILE REFERENCE: 35052/204375
; CURRENT APPLICATION NUMBER: US/10/681,970
; CURRENT FILING DATE: 2003-10-09

Db 1395 SFTPVNSLDPLLLTRYLRHQPQSWHQAIALRMEVLGCEAODLY 1438

RESULT 2
US-10-047-257-1
; Sequence 1, Application US/10047257
; Publication No. US2002011512A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255.2
; CURRENT APPLICATION NUMBER: US/10/047,257
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-047-257-1

Query Match 100.0%; Score 871; DB 13; Length 1438;
Best Local Similarity 100.0%; Pred. No. 2.4e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDLNSCSMPLGMSKAIISDAQITASSYFTNMFTWSPSKARLHLOGRSNAWRPQVNNPKE 60
Db 1275 CDLNSCSMPLGMSKAIISDAQITASSYFTNMFTWSPSKARLHLOGRSNAWRPQVNNPKE 1334

Qy 61 WLQVDFQKTMKVGTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 1335 WLQVDFQKTMKVGTGTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 1394

Qy 121 SFTPVNSLDPLLLTRYLRHQPQSWHQAIALRMEVLGCEAODLY 164
Db 1395 SFTPVNSLDPLLLTRYLRHQPQSWHQAIALRMEVLGCEAODLY 1438

RESULT 3
US-10-225-900-1
; Sequence 1, Application US/10225900
; Publication No. US2003007752A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255
; CURRENT APPLICATION NUMBER: US/10/225,900
; CURRENT FILING DATE: 2002-08-22
; PRIOR APPLICATION NUMBER: US/09/209,916
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-225-900-1

Query Match 100.0%; Score 871; DB 14; Length 1438;
Best Local Similarity 100.0%; Pred. No. 2.4e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

;; PRIOR APPLICATION NUMBER: US/09/689,430
;; PRIOR FILING DATE: 2001-08-22
;; PRIOR APPLICATION NUMBER: 60/158,780
;; PRIOR FILING DATE: 1999-10-12
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 1471
;; TYPE: PRT
;; ORGANISM: Homo sapiens B-domain deleted factor VIII
;; FEATURE:
;; OTHER INFORMATION: Homo sapiens BDD FVIII
US-10-661-970-2

Query Match 100.0%; Score 871; DB 12; Length 1471;
Best Local Similarity 100.0%; Pred. No. 2.5e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCMPLGWSKASDAQITASSYFTNMFATWSPSKARHLQGRSNARPPQVNNPKE 60
Db 1308 CDLNSCMPLGWSKASDAQITASSYFTNMFATWSPSKARHLQGRSNARPPQVNNPKE 1367
QY 61 WLQVDFQKTMKVTGTTGGVKSLLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 1368 WLQVDFQKTMKVTGTTGGVKSLLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQD 1427
QY 121 SFTPVNSLDPPLLTRYLRHPOSVWHQIALRMEVLGCEAODLY 164
Db 1428 SFTPVNSLDPPLLTRYLRHPOSVWHQIALRMEVLGCEAODLY 1471

RESULT 6
US-10-095-718-2
;; Sequence 2, Application US/10095718
;; Publication No. US20020131956A1
;; GENERAL INFORMATION:
;; APPLICANT: Walsh, Christopher
;; APPLICANT: Chao, Hengjun
;; APPLICANT: Burstein, Haim
;; APPLICANT: Lynch, Carmel
;; APPLICANT: Stepan, Tony
;; APPLICANT: Munson, Keith
;; TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
;; TITLE OF INVENTION: Methods of Using the Same
;; FILE REFERENCE: 35052/204375
;; CURRENT APPLICATION NUMBER: US/10/095,718
;; PRIOR FILING DATE: 2002-03-12
;; PRIOR APPLICATION NUMBER: 09/689,430
;; PRIOR FILING DATE: 2001-08-22
;; PRIOR APPLICATION NUMBER: 60/158,780
;; PRIOR FILING DATE: 1999-10-12
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 1471
;; TYPE: PRT
;; ORGANISM: Homo sapiens B-domain deleted factor VIII
;; FEATURE:
;; OTHER INFORMATION: Homo sapiens BDD FVIII
US-10-095-718-2

Query Match 100.0%; Score 871; DB 13; Length 1471;
Best Local Similarity 100.0%; Pred. No. 2.5e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCMPLGWSKASDAQITASSYFTNMFATWSPSKARHLQGRSNARPPQVNNPKE 60
Db 1308 CDLNSCMPLGWSKASDAQITASSYFTNMFATWSPSKARHLQGRSNARPPQVNNPKE 1367
QY 61 WLQVDFQKTMKVTGTTGGVKSLLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 1368 WLQVDFQKTMKVTGTTGGVKSLLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQD 1427

QY 121 SFTPVNSLDPPLLTRYLRHPOSVWHQIALRMEVLGCEAODLY 164
Db 1428 SFTPVNSLDPPLLTRYLRHPOSVWHQIALRMEVLGCEAODLY 1471

RESULT 7
US-09-957-641-2
;; Sequence 2, Application US/09957641
;; Publication No. US20020182670A1
;; GENERAL INFORMATION:
;; APPLICANT: Emory University
;; TITLE OF INVENTION: MODIFIED FACTOR VIII
;; FILE REFERENCE: 75-00
;; CURRENT APPLICATION NUMBER: US/09/957,641
;; CURRENT FILING DATE: 2001-09-16
;; PRIOR APPLICATION NUMBER: US 60/234047
;; PRIOR FILING DATE: 2000-09-19
;; PRIOR APPLICATION NUMBER: US 60/236460
;; PRIOR FILING DATE: 2000-09-29
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 2
;; LENGTH: 2332
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-957-641-2

Query Match 100.0%; Score 871; DB 9; Length 2332;
Best Local Similarity 100.0%; Pred. No. 4.6e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCMPLGWSKASDAQITASSYFTNMFATWSPSKARHLQGRSNARPPQVNNPKE 60
Db 2169 CDLNSCMPLGWSKASDAQITASSYFTNMFATWSPSKARHLQGRSNARPPQVNNPKE 2228
QY 61 WLQVDFQKTMKVTGTTGGVKSLLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 2229 WLQVDFQKTMKVTGTTGGVKSLLTSMYKKEFLISSQDGHQWTLFFQNGKVKVFGNQD 2288
QY 121 SFTPVNSLDPPLLTRYLRHPOSVWHQIALRMEVLGCEAODLY 164
Db 2289 SFTPVNSLDPPLLTRYLRHPOSVWHQIALRMEVLGCEAODLY 2332

RESULT 8
US-10-187-319-2
;; Sequence 2, Application US/10187319
;; Publication No. US20030068785A1
;; GENERAL INFORMATION:
;; APPLICANT: Lollar, John S.
;; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
;; NUMBER OF SEQUENCES: 40
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
;; STREET: 5370 Manhattan Circle Suite 201
;; CITY: Boulder
;; STATE: Colorado
;; COUNTRY: USA
;; ZIP: 80303
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION NUMBER: US/10/187,319
;; FILING DATE: 27-Aug-2002
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 09/523,656
;; FILING DATE: 2000-03-10
;; APPLICATION NUMBER: US 09/037,601
;; FILING DATE: 1998-03-10

```

; APPLICATION NUMBER: WO PCT/US97/11155
; FILING DATE: 1997-06-26
; APPLICATION NUMBER: US 08/670,707
; FILING DATE: 1996-06-26
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorance L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: <Unknown>
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-187-319-2

Query Match      100.0%; Score 871; DB 14; Length 2332;
Best Local Similarity 100.0%; Pred. No. 4.6e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CDLNSCSMPLGMSKATSDAQITASSYFTNMFWATSPSKARLHLOGRSNARPPQVNNPKE 60
DB      2169 CDLNSCSMPLGMSKATSDAQITASSYFTNMFWATSPSKARLHLOGRSNARPPQVNNPKE 2228

QY      61 WLQVDFQKTMKVGTGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
DB      2229 WLQVDFQKTMKVGTGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 2288

QY      121 SFTPVNSLDPPLTRYLRHPQSWVHQIALRMEVLGCEAODLY 164
DB      2289 SFTPVNSLDPPLTRYLRHPQSWVHQIALRMEVLGCEAODLY 2332

RESULT 9
US-10-131-510A-2
; Sequence 2, Application US/10131510A
; Publication No. US20030166536A1
; GENERAL INFORMATION:
; APPLICANT: Lollat, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95J
; CURRENT APPLICATION NUMBER: US/10/131,510A
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: U.S. 09/315,179
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: U.S. 09/037,601
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: U.S. 08/670,707
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: PCT/US97/11155
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: PCT/US94/13200
; PRIOR FILING DATE: 1994-11-15
; PRIOR APPLICATION NUMBER: U.S. 08/212,133
; PRIOR FILING DATE: 1994-03-11
; PRIOR APPLICATION NUMBER: U.S. 07/864,004
; PRIOR FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
US-10-131-510A-2

Query Match      100.0%; Score 871; DB 14; Length 2332;
Best Local Similarity 100.0%; Pred. No. 4.6e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CDLNSCSMPLGMSKATSDAQITASSYFTNMFWATSPSKARLHLOGRSNARPPQVNNPKE 60
DB      2169 CDLNSCSMPLGMSKATSDAQITASSYFTNMFWATSPSKARLHLOGRSNARPPQVNNPKE 2228

QY      61 WLQVDFQKTMKVGTGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
DB      2229 WLQVDFQKTMKVGTGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 2288

QY      121 SFTPVNSLDPPLTRYLRHPQSWVHQIALRMEVLGCEAODLY 164
DB      2289 SFTPVNSLDPPLTRYLRHPQSWVHQIALRMEVLGCEAODLY 2332

RESULT 10
US-10-445-235-2
; Sequence 2, Application US/10445235
; Publication No. US20040005670A1
; GENERAL INFORMATION:
; APPLICANT: Katherine A. High
; APPLICANT: Rodney M. Camire
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT OF HEMOPHILIA A
; FILE REFERENCE: CHOP0176
; CURRENT APPLICATION NUMBER: US/10/445,235
; CURRENT FILING DATE: 2003-05-22
; PRIOR APPLICATION NUMBER: 60/382,486
; PRIOR FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-445-235-2

Query Match      100.0%; Score 871; DB 15; Length 2332;
Best Local Similarity 100.0%; Pred. No. 4.6e-90;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CDLNSCSMPLGMSKATSDAQITASSYFTNMFWATSPSKARLHLOGRSNARPPQVNNPKE 60
DB      2169 CDLNSCSMPLGMSKATSDAQITASSYFTNMFWATSPSKARLHLOGRSNARPPQVNNPKE 2228

QY      61 WLQVDFQKTMKVGTGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
DB      2229 WLQVDFQKTMKVGTGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 2288

QY      121 SFTPVNSLDPPLTRYLRHPQSWVHQIALRMEVLGCEAODLY 164
DB      2289 SFTPVNSLDPPLTRYLRHPQSWVHQIALRMEVLGCEAODLY 2332

RESULT 11
US-10-360-101-229
; Sequence 229, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309

```

```

; SOFTWARE: Patentin version 3.1
; SEQ ID NO: 229
; LENGTH: 232
; TYPE: prt
; ORGANISM: Artificial Sequence
; FEATURE:
;   - OTHER INFORMATION: sequence of factor VIII
US-10-1360-101-229

```

	Query Match	Best Local Similarity	Matches	100.0%;	Score	871;	DB	15;	Length	2332;
	Matches	164;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	CDLNSCSMP	LGMSKAISDAQI	TASSYFTNNW	TWSPSKARLHLQGRSNAWRPQVNNPKE	60				
Db	2169	CDLNSCSMP	LGMSKAISDAQI	TASSYFTNNF	ATWSPSKARLHLQGRSNAWRPQVNNPKE	2228				
Qy	61	WLQVDFQK	TKWKTGVT	TQGVKSL	LTSMYVKEFLISSQDGHQMTLFFQNGKVKVFOGQND	120				
Db	2229	WLQVDFQK	TKWKTGVT	TQGVKSL	LTSMYVKEFLISSQDGHQMTLFFQNGKVKVFOGQND	2288				
Qy	121	SFTFPVNSL	DPPLTRVLR	LTHPSQSWH	QIARLMEVLGCEAQDLY	164				
Db	2289	SFTFPVNSL	DPPLTRVLR	LTHPSQSWH	QIARLMEVLGCEAQDLY	2332				

```

RESULT 12
US - 10-239-498A-2
; Sequence 2, Application US/10239498A
; Publication No. US2004002333A1
; GENERAL INFORMATION:
; APPLICANT: Hauser, Charlotte
; APPLICANT: Horster, Andrea
; APPLICANT: Schroder, Carola
; APPLICANT: Lehneier, Michael
; TITLE OF INVENTION: Production of R
; TITLE OF INVENTION: Human Cell Lin
; FILE REFERENCE: 80977.0001
; CURRENT APPLICATION NUMBER: US/10/2
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: PCT/EP01/
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US - 10-239-498A-2

```

Query Match	100.0%	Score 871;	DB 16;	Length 2332;
Best Local Similarity	100.0%;	Pred. No. 4.6e-90;		
Matches 164;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CDLNSCSPLGMSKAI	SDAQITASSYFTNNFATWSPSKALH	LHQGSNAWRPOVNNPKE 60
Db	2169	CDLNSCSPLGMSKAI	SDAQITASSYFTNNFATWSPSKALH	LHQGSNAWRPOVNNPKE 2228
QY	61	WLQVDFQKTMKVGT	YTTGCVKSLTSMYVK	EFLTSSSDGQHWTLTFQNGKVKVFGNQD 120
Db	2229	WLQVDFQKTMKVGT	YTTGCVKSLTSMYVK	EFLTSSSDGQHWTLTFQNGKVKVFGNQD 2288
QY	121	SFTPVNSLDPLLT	RYLRHHQSVHQAIALRMEV	LGC EAQDLY 164
Db	2289	SFTPVNSLDPLLT	RYLRHHQSVHQAIALRMEV	LGC EAQDLY 2332

RESULT 13
US-10-466-998A-1
; Sequence 1, Application US/10466998A
; Publication No. US20040126936A1
; GENERAL INFORMATION:
; APPLICANT: BAJAJ, S. Paul

```

; APPLICANT: FAY, Philip J.
; TITLE OF INVENTION: Factor IXa: Factor VIIIa Interaction and Methods
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: 66153-41436
; CURRENT APPLICATION NUMBER: US/10/466,998A
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: PCT/US02/01724
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US 60/263,431
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-466-998A-1

```

	Query Match	100.0.0%;	Score 871;	DB 16;	Length 2332;
	Best local Similarity	100.0%;	Pred. No. 4.6e-90;		
	Matches 164;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CDLNSCSMPGMSKATSDAQIITASSYFTNNKFTWSPSKARLHLQGSNAWRPQVNNPKE	60		
Db	2169	CDLNSCSMPGMSKATSDAQIITASSYFTNNKFTWSPSKARLHLQGSNAWRPQVNNPKE	2238		
Qy	61	WLQVDFQKTKWVTGTTTCGVKSLTSMYVKEFLISSQDGHQWTLFPQNGKVKVFGQNQD	120		
Db	2229	WLQVDFQKTKWVTGTTTCGVKSLTSMYVKEFLISSQDGHQWTLFPQNGKVKVFGQNQD	2288		
Qy	121	SFTPEVNSLDPPLLTRYLRIHPQSWHQAIALRMEVLGCEAODLY	164		
Db	2289	SFTPEVNSLDPPLLTRYLRIHPQSWHQAIALRMEVLGCEAODLY	2332		

RESULT 14
US-10-411-037-30
; Sequence 30, Application US/10411037
; Publication No. US2004003446A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: Defrees, Shawn
; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bose, Caryn
; TITLE OF INVENTION: ALPHA GALACTOSIDASE A
; TITLE OF INVENTION: GALACTOSIDASE A
; FILE REFERENCE: 040853-01-5082
; CURRENT APPLICATION NUMBER: US/10/411,037
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-037-30

Search completed: September 29, 2004, 16:51:38
Job time : 132 secs

	Query Match	100.0%;	Score 871;	DB 12;	Length 2351;
	Best Local Similarity	100.0%;	Pred. No. 4.7e-90;		
	Matches 164;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CDLNSCMPGMSKALSDAQIITASSYFTNMETWSPSKARLHLQGRSNARPOVNNPKE	60		
Db	2188	CDLNSCMPGMSKALSDAQIITASSYFTNMETWSPSKARLHLQGRSNARPOVNNPKE	2247		
Qy	61	WLQVDFQTKMKVTGVTITQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNQD	120		
Db	2248	WLQVDFQTKMKVTGVTITQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGQNQD	2307		
Qy	121	SFTFPVNSLDPPLLTYLRLTHPSQWHTQIALRNEVLGCEAQDLY	164		
Db	2308	SFTFPVNSLDPPLLTYLRLTHPSQWHTQIALRNEVLGCEAQDLY	2351		

RESULT 15
US-10-411-026-30
; Sequence 30, Application US/10411026
; Publication No. US20040063911A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DeFrees, Shawn
; APPLICANT: Zepf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; TITLE OF INVENTION: PROTEIN REMODELING METHODS AND PEPTIDES PRODUCED BY THE
; FILE REFERENCE: METHODS
; FILE REFERENCE: 040853-01-5053
; CURRENT APPLICATION NUMBER: US/10/411,026
; CURRENT FILING DATE: 2003-04-09
; PRIOR APPLICATION NUMBER: US 60/328,523
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/344,692
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/387,292
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 60/391,777
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/396,594
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: US 60/404,249
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 2351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-411-026-30

Query Match	100.0%;	Score 871;	DB 12;	Length 2351;
Best Local Similarity	100.0%;	Pred. No. 4.7e-90;		
Matches 164;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CDLNSCSNPLGMSKAI	SDAQITASSYFTNNYFATWSPSKARHLHQR	SNAWRPVNNPK 60
Ddb	2188	CDLNSCSNPLGMSKAI	SDAQITASSYFTNNYFATWSPSKARHLHQR	SNAWRPVNNPK 2247
QY	61	WLQVDFQTKMKVGT	VTGQVKSLITSMYKFEFLISSODGQWTL	TFPONGKVKVFGNQD 120
Ddb	2248	WLQVDFQTKMKVGT	VTGQVKSLITSMYKFEFLISSODGQWTL	TFPONGKVKVFGNQD 2307
QY	121	SFTFVNSLDPPL	TLRYLRIHQSWHQIALRMEVLGCEA	QADLY 164
Ddb	2308	SFTFVNSLDPPL	TLRYLRIHQSWHQIALRMEVLGCEA	QADLY 2351

34	184	21.1	1284	1	NRX4_DROME
35	183	21.0	910	1	DDR1_RAT
36	183	21.0	911	1	DDR1_MOUSE
37	181	20.8	1384	1	CTAL_HUMAN
38	180.5	20.7	854	1	DDR2_MOUSE
39	176.5	20.3	855	1	DDR2_HUMAN
40	171	19.6	1381	1	CTAL_RAT
41	171	19.6	1385	1	CTAL_MOUSE
42	107.5	12.3	257	1	DIS2_DICDI
43	104	11.9	867	1	SSPO_BOVIN
44	84.5	9.7	149	1	DISD_DICDI
45	84.5	9.7	253	1	DISC_DICDI

ALIGNMENTS

RESULT 1					
FA8_HUMAN	FA8_HUMAN	STANDARD;	PRT;	2351	AA.
ID	AC	P00451;			
DT	21-JUL-1986	(Rel. 01, Created)			
DT	21-JUL-1986	(Rel. 01, Last sequence update)			
DT	10-OCT-2003	(Rel. 42, Last annotation update)			
DE	Coagulation factor VIII precursor (Procoagulant component)				
DE	(Antihemophilic factor) (AHF).				
DE	F8 OR F8C.				
GN	F8 OR F8C.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
OX	[1]				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=86081164; PubMed=3935400;				
RA	Truett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,				
RA	Hartog K., Kuo C.H., Maslary F.R., Merryweather J.P., Najarian R.,				
RA	Pachl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,				
RA	Urdea M.S., Valenzuela P., Dahl H.H.M., Favalaro J., Hansen J.,				
RA	Nordfang O., Ezban M.				
RT	"Characterization of the polypeptide composition of human factor				
RT	VIII:C and the nucleotide sequence and expression of the human kidney				
RT	cDNA."				
RL	DNA 4:333-349(1985).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=85061548; PubMed=6438526;				
RA	Wood W.I., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J.,				
RA	Keyt B., Seeburg P.H., Smith D.H., Hollingshead P., Wion K.L.,				
RA	Delwart E., Tuddenham E.G.D., Vehar G.A., Lawn R.M.;				
RA	"Expression of active human factor VIII from recombinant DNA clones."				
RL	Nature 312:330-337(1984).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=85061550; PubMed=6438528;				
RA	Toole J.J., Knopf J.L., Wozney J.M., Sultzman L.A., Buecker J.L.,				
RA	Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,				
RA	Amchelt G.W., Foster W.B., Coe M.L., Knutson G.J., Pass D.N.,				
RA	Hewick R.M.;				
RT	"Molecular cloning of a cDNA encoding human antihaemophilic factor."				
RT	Nature 312:342-347(1984).				
RL	[4]				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=93265012; PubMed=1303178;				
RA	Gitschier J., Wood W.I.;				
RA	"Sequence of the exon-containing regions of the human factor VIII				
RT	gene."				
RL	Hum. Mol. Genet. 1:199-200(1992).				
RN	[5]				
RP	SEQUENCE OF 2064-2070 FROM N.A.				
RA	de Water N.S., Williams R., Browett P.J.;				
RA	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.				
RL	[6]				
RP	SULFATION OF TYR-1699.				

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - protein search, using sw model
Run on: September 29, 2004, 16:35:13 ; Search time 13 seconds
(without alignments)
656.885 Million cell updates/sec
Title: US-10-049-399A-1_COPY_2169_2332
Perfect score: 871
Sequence: 1 CDLNSCMPLGHEKSAISDA.....VHQIALRMENVLCGCAQDLY 164
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	ID Description
1	871	100.0	1 FA8_HUMAN
2	752	86.3	2319 1 FA8_MOUSE
3	720	82.7	2133 1 FA8_PIG
4	385	44.2	2258 1 FA5_PIG
5	379	43.5	2211 1 FA5_BOVIN
6	376	43.2	2224 1 FA5_HUMAN
7	367.5	42.2	427 1 MFGM_MOUSE
8	353.5	40.6	463 1 MFGM_MOUSE
9	329	37.8	925 1 NRP2_RAT
10	328.5	37.7	409 1 MFGM_PIG
11	327.5	37.6	387 1 MFGM_HUMAN
12	327	37.5	931 1 NRP2_HUMAN
13	327	37.5	931 1 NRP2_MOUSE
14	324.5	37.3	427 1 MFGM_BOVIN
15	275.5	31.6	914 1 NRP1_CHICK
16	269.5	30.9	923 1 NRP1_MOUSE
17	266	30.5	764 1 CPX2_MOUSE
18	265.5	30.5	922 1 NRP1_RAT
19	261	30.5	923 1 NRP1_HUMAN
20	261	30.0	756 1 CPX2_HUMAN
21	256	29.4	280 1 XLR1_FUGRU
22	254.5	29.2	928 1 NRP1_XENLA
23	247	28.4	224 1 XLR1_MOUSE
24	243	27.9	224 1 XLR1_HUMAN
25	234.5	26.9	3133 1 HMCT_BOMMO
26	217	24.9	1331 1 CTAT2_HUMAN
27	215.5	24.7	722 1 CPXM_MOUSE
28	214	24.6	734 1 CPXM_HUMAN
29	208	23.9	1310 1 CTAT4_MOUSE
30	206	23.7	1308 1 CTAT4_HUMAN
31	199	22.8	909 1 DDR1_PANTR
32	194	22.3	1288 1 CTAT3_HUMAN
33	192	22.0	913 1 DDR1_HUMAN

RX MEDLINE=91093266; PubMed=1898735;
 RA Leyte A., van Schijndel H.B., Niehrs C., Huttner W.B., Verbeet M.P.,
 RA Mertens K., van Mourik J.A.;
 RT "Sulfation of Tyr1680 of human blood coagulation factor VIII is
 RT essential for the interaction of factor VIII with von Willebrand
 RT factor.";
 RL J. Biol. Chem. 266:740-746(1991).
 RN [7]
 RP VARIANT HEMA GYL-291.
 RX MEDLINE=92207952; PubMed=1554716;
 RA Pitman D.D., Wang J.H., Kaufman R.J.;
 RT "Identification and functional importance of tyrosine sulfate
 RT residues within recombinant factor VIII.";
 RL Biochemistry 31:3315-3325(1992).
 RN [8]
 RP STRUCTURE BY NMR OF 2322-2343.
 RX MEDLINE=95200924; PubMed=7893714;
 RA Gilbert G.E., Bateja J.D.;
 RT "Membrane-binding peptide from the C2 domain of factor VIII forms an
 RT amphipathic structure as determined by NMR spectroscopy.";
 RL Biochemistry 34:3022-3031(1995).
 RN [9]
 RP REVIEW ON MOLECULAR BASIS OF HEMA.
 RX MEDLINE=91221499; PubMed=1902642;
 RA Gitschier J.;
 RT "The molecular basis of hemophilia A.";
 RL Ann. N.Y. Acad. Sci. 614:89-96(1991).
 RN [10]
 RP REVIEW ON MOLECULAR BASIS OF HEMA.
 RX MEDLINE=89088506; PubMed=2491949;
 RA White G.C. II, Shoemaker C.B.;
 RT "Factor VIII gene and hemophilia A.";
 RL Blood 73:1-12(1989).
 RN [11]
 RP REVIEW ON MOLECULAR BASIS OF HEMA.
 RX MEDLINE=95245332; PubMed=7728145;
 RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
 RT "Molecular etiology of factor VIII deficiency in hemophilia A.";
 RL Hum. Mutat. 5:1-22(1995).
 RN [12]
 RP VARIANT HEMA GLN-2326.
 RX MEDLINE=86235434; PubMed=3012775;
 RA Gitschier J., Wood W.I., Shuman M.A., Lawn R.M.;
 RT "Identification of a missense mutation in the factor VIII gene of a
 RT mild hemophiliac.";
 RL Science 232:1415-1416(1986).
 RN [13]
 RP VARIANT HEMA PRO-2135.
 RX MEDLINE=88096539; PubMed=3122181;
 RA Levinson B., Janco R.L., Phillips J.A. III, Gitschier J.;
 RT "A novel missense mutation in the factor VIII gene identified by
 RT analysis of amplified hemophilia DNA sequences.";
 RL Nucleic Acids Res. 15:9797-9805(1987).
 RN [14]
 RP VARIANT HEMA GLN-2228.
 RX MEDLINE=88191889; PubMed=2833855;
 RA Youssoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
 RA Kazazian H.H.;
 RT "Nonsense and missense mutations in hemophilia A: estimate of the
 RT relative mutation rate at CG dinucleotides.";
 RL Am. J. Hum. Genet. 42:718-725(1988).
 RN [15]
 RP VARIANT HEMA GYL-291.
 RX MEDLINE=88220354; PubMed=2835904;
 RA Youssoufian H., Wong C., Aronis S., Platakoukis H., Kazazian H.H. Jr.,
 RA Antonarakis S.E.;
 RT "Moderately severe hemophilia A resulting from Glu-->Gly substitution
 RT in exon 7 of the factor VIII gene.";
 RL Am. J. Hum. Genet. 42:867-871(1988).
 RN [16]
 RP VARIANT HEMA GYL-291.
 RX MEDLINE=89274393; PubMed=2499363;
 RA O'Brien D.P., Tuddenham E.G.;
 RT "Purification and characterization of factor VIII 1,689-Cys: a
 RT nonfunctional cofactor occurring in a patient with severe hemophilia
 RT A.";
 RL Blood 73:2117-2122(1989).
 RN [17]
 RP VARIANT HEMA GYL-291.
 RX MEDLINE=90001543; PubMed=2506948;
 RA Shima M., Ware J., Yoshioka A., Fukui H., Fulcher C.A.;
 RT "An arginine to cysteine amino acid substitution at a critical
 RT thrombin cleavage site in a dysfunctional factor VIII molecule.";
 RL Blood 74:1612-1617(1989).
 RN [18]
 RP VARIANT HEMA LEU-189.
 RX MEDLINE=90057680; PubMed=2510835;
 RA Chan V., Chan T.K., Tong T.M., Todd D.;
 RT "A novel missense mutation in exon 4 of the factor VIII:C gene
 RT resulting in moderately severe hemophilia A.";
 RL Blood 74:2688-2691(1989).
 RN [19]
 RP VARIANT HEMA LEU-2326.
 RX MEDLINE=89197216; PubMed=2495245;
 RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
 RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
 RT of the factor VIII gene.";
 RL Hum. Genet. 81:335-338(1989).
 RN [20]
 RP VARIANT HEMA HIS-391.
 RX MEDLINE=89264602; PubMed=2498882;
 RA Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Fujimaki M., Hoyer L.W.;
 RT "Direct characterization of factor VIII in plasma: detection of a
 RT mutation altering a thrombin cleavage site
 RT (arginine-372-->histidine).";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
 RN [21]
 RP VARIANT HEMA GYL-291.
 RX MEDLINE=90105723; PubMed=2104766;
 RA Arai M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Phillips J.A. III, Janco R.L., Hoyer L.W.;
 RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to
 RT Cys) in the factor VIII gene of two unrelated patients with
 RT cross-reacting material-positive hemophilia A.";
 RL Blood 75:384-389(1990).
 RN [22]
 RP VARIANTS HEMA GLN-2228 AND LEU-2326.
 RX MEDLINE=90123183; PubMed=2105106;
 RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
 RA Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;
 RT "Recurrent mutations and three novel rearrangements in the factor
 RT VIII gene of hemophilia A patients of Italian descent.";
 RL Blood 75:662-670(1990).
 RN [23]
 RP VARIANT HEMA GYL-291.
 RX MEDLINE=90329422; PubMed=1973901;
 RA Pattinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
 RT "CRM+ hemophilia A due to a missense mutation (372-->Cys) at the
 RT internal heavy chain thrombin cleavage site.";
 RL Br. J. Haematol. 75:73-77(1990).
 RN [24]
 RP VARIANTS HEMA PHE-1699 AND GYL-1708.
 RX MEDLINE=90152691; PubMed=2105906;
 RA Higuchi M., Wong C., Kochhan L., Olek K., Aronis S., Kasper C.K.,
 RA Kazazian H.H., Antonarakis S.E.;
 RT "Characterization of mutations in the factor VIII gene by direct
 RT sequencing of amplified genomic DNA.";
 RL Genomics 6:65-71(1990).
 RN [25]
 RP VARIANTS HEMA GYL-291 AND ASP-1941.
 RX MEDLINE=90169988; PubMed=2106480;
 RA Traystman M.D., Higuchi M., Kasper C.K., Antonarakis S.E.,
 RA Kazazian H.H.;
 RT "Use of denaturing gradient gel electrophoresis to detect point
 RT mutations in the factor VIII gene."

FT	CHAIN	20	2319	COAGULATION FACTOR VIII.
FT	DOMAIN	20	349	F5/8 TYPE A 1.
FT	DOMAIN	20	199	PLASTOCYANIN-LIKE 1.
FT	DOMAIN	207	349	PLASTOCYANIN-LIKE 2.
FT	DOMAIN	399	730	F5/8 TYPE A 2.
FT	DOMAIN	399	573	PLASTOCYANIN-LIKE 3.
FT	DOMAIN	583	730	PLASTOCYANIN-LIKE 4.
FT	DOMAIN	760	1640	B.
FT	DOMAIN	1683	2008	F5/8 TYPE A 3.
FT	DOMAIN	1683	1845	PLASTOCYANIN-LIKE 5.
FT	DOMAIN	1855	2008	PLASTOCYANIN-LIKE 6.
FT	DOMAIN	2008	2156	F5/8 TYPE C 1.
FT	DOMAIN	2161	2313	F5/8 TYPE C 2.
FT	SITE	391	392	CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT	SITE	759	760	CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT	SITE	1678	1679	CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT	SITE	1324	1325	CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT	SITE	1640	1641	CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT	MOD_RES	367	367	SULFATION (BY SIMILARITY).
FT	MOD_RES	737	737	SULFATION (BY SIMILARITY).
FT	MOD_RES	738	738	SULFATION (BY SIMILARITY).
FT	MOD_RES	742	742	SULFATION (BY SIMILARITY).
FT	MOD_RES	1669	1669	SULFATION (REQUIRED FOR VWF BINDING) (BY SIMILARITY).
FT	MOD_RES	1687	1687	SULFATION (BY SIMILARITY).
FT	DISULFID	173	199	PROBABLE.
FT	DISULFID	547	573	PROBABLE.
FT	DISULFID	1819	1845	PROBABLE.
FT	DISULFID	2008	2156	BY SIMILARITY.
FT	DISULFID	2161	2313	BY SIMILARITY.
FT	CARBOHYD	61	61	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	233	233	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	259	259	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	423	423	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	601	601	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	880	880	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	958	958	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1015	1015	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1022	1022	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1026	1026	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1044	1044	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1076	1076	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1087	1087	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1136	1136	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1161	1161	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1192	1192	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1255	1255	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1368	1268	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1373	1273	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1274	1274	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1302	1302	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1316	1316	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1340	1340	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1378	1378	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1797	1797	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	2105	2105	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	SEQUENCE	2319	AA; FD054DE051DB2A01 CRC64;	
QY	Match	86.3%;	Score 752; DB 1; Length 2319;	
Db	Best Local Similarity	84.1%;	Pred. No. 1.3e-64;	
QY	Matches 138; Conservative	11;	Mismatches 15; Indels 0; Gaps 0;	
QY	1	CDLNSCSPIPLGMSKAISSDAQITASSYFTNFWATSPSKARLHLQGRSNARWQVNNPK 60		
Db	2156	CDLNSCSPIPLGMSKVIISDTQITASSYFTNFWATSPSQARLHLQGRTNARWQVNDPKQ 2215		
QY	61	WLQVDFOKTMKVTTGGVKSLLTSMYKVEFLISSODGHQWTLFPQNGKVKVQGNQD 120		
Db	2216	WLQVDLQKTMKVTTGIITGGVKSLLTSMYKVEFLISSODGHHTQIILYNGKVKVQGNQD 2275		
QY	121	SFTPVNSLDPLLTRYLRLRHHPQSWHQAIRMEVLGCEAQDLY 164		
Db	2276	SFTPVNSLDPLLTRYLRLRHHPQSWHQAIRMEVLGCEAQDLY 2319		

Query Match	86.3%	Score	752	DB 1	Length	2319
Best Local Similarity	84.1%	Pred. No.	1.3e-64			
Matches	138	Conservative	11	Mismatches	15	Indels
					0	Gaps
0						
QY	1	CDLNSCSPLGNESKAISSDAQITASSYFTNNFATWSPSKARHLQGRSNARWRPQVNNPKF	60			
DB	2156	CDLNSCS:PLGNESKVIDTQITASSYFTNNFATWSPQARHLQGRTNARWPQVNDPKQ	2215			
QY	61	WLQVDFOKTMKVGTITGGVKSLTSMYVKEFLISSQDGHQHTLFPQNGKVKVQGNOD	120			
DB	2216	WLQVDLQKTMKVGTITGGVKSLFTSMFVKEFLISSQDGHHTQTILYNGKVKVQGNOD	2275			
QY	121	SFTPVNSLDPLLTRYLRIHPQSWVHCIALRMEVLGCEAQDLY	164			
DB	2276	SFTPVNSLDPLLTRYLRIHPQIWEHQIALRLIILGCEAQQY	2319			

RESULT 3
FAS_PIG STANDARD; PRT; 2133 AA.
AC P12763; Q95243; PRT; 2133 AA.
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Coagulation factor VIII precursor (Procoagulant component).
GN F8 OR C8.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Healey J.F., Lubin I.M., Lollar P.,
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 705-1573 FROM N.A.
RX MEDLINE=86287369; PubMed=3016730;
RA Toole J.J., Pittman D.D., Orr E.C., Murtha P., Wasley L.C.,
RA Kaufman R.J.;
RT "A large region (approximately equal to 95 kDa) of human factor VIII
is dispensable for in vitro procoagulant activity.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5939-5942 (1986).
RN [3]
RP SEQUENCE OF 392-759 FROM N.A.
RX MEDLINE=94179260; PubMed=7510693;
RA Lubin I.M., Healey J.F., Scandella D., Runge M.S., Lollar P.;
RT "Elimination of a major inhibitor epitope in factor VIII.";
RL J. Biol. Chem. 269:8638-8641 (1994).
CC -!- FUNCTION: Factor VIII, along with calcium and phospholipid, acts
CC as a cofactor for factor IXa when it converts factor X to the
CC activated form, factor Xa.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: Contains 3 F5/8 type A domains.
CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC -!- SIMILARITY: STRONG, TO COAGULATION FACTOR V.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U49517; AAB06705.1; -.
DR PIR; A25945; A25945.
DR PIR; T42763; T42763.
DR HSP; P00451; 1CFG.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR000421; FAS8 C.
DR InterPro; IPR008979; Gal Bind-like.
DR Pfam; PF00394; Cu-oxidase; 3.
DR Pfam; PF00754; F5 F8 type C; 2.
DR SMART; SM00231; FAS8G; 2.
DR PROSITE; PS01285; FAS6C_1; 2.
DR PROSITE; PS01286; FAS6C_2; 2.
DR PROSITE; PS00022; FAS8C_3; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
KW Blood coagulation; Repeat; Plasma; Acute phase; Calcium;
KW Signal; Glycoprotein; Sulfation.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 2133 COAGULATION FACTOR VIII.
FT DOMAIN 20 357 F5/8 TYPE A 1.
FT DOMAIN 20 199 PLASTOCYANIN-LIKE 1.
FT DOMAIN 207 357 PLASTOCYANIN-LIKE 2.
FT DOMAIN 399 730 F5/8 TYPE A 2.
FT DOMAIN 399 573 PLASTOCYANIN-LIKE 3.

FT DOMAIN 583 730 PLASTOCYANIN-LIKE 4.
FT DOMAIN 760 1599 B.
FT DOMAIN 1495 1822 F5/8 TYPE A 3.
FT DOMAIN 1495 1659 PLASTOCYANIN-LIKE 5.
FT DOMAIN 1669 1822 PLASTOCYANIN-LIKE 6.
FT DOMAIN 1822 1970 F5/8 TYPE C 1.
FT DOMAIN 1975 2127 F5/8 TYPE C 2.
FT SITE 391 392 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 759 760 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT SITE 1449 1450 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT SITE 1490 1491 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT MOD_RES 737 737 SULFATION (BY SIMILARITY).
FT MOD_RES 738 738 SULFATION (BY SIMILARITY).
FT MOD_RES 742 742 SULFATION (BY SIMILARITY).
FT DISULFID 173 199 PROBABLE.
FT DISULFID 547 573 PROBABLE.
FT DISULFID 1633 1659 PROBABLE.
FT DISULFID 1822 1970 BY SIMILARITY.
FT DISULFID 1975 2127 BY SIMILARITY.
FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 929 929 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 985 985 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1025 1025 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1111 1111 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1181 1181 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1208 1208 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1245 1245 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1265 1265 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1335 1335 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1408 1408 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1611 1611 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1919 1919 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 713 713 N -> M (IN REF. 2).
FT CONFLICT 734 734 I -> T (IN REF. 2).
FT CONFLICT 792 792 G -> Q (IN REF. 2).
FT CONFLICT 1133 1133 E -> F (IN REF. 2).
FT CONFLICT 1191 1191 I -> L (IN REF. 2).
FT CONFLICT 1209 1209 R -> F (IN REF. 2).
FT CONFLICT 1437 1437 C -> G (IN REF. 2).
FT CONFLICT 1456 1456 F -> R (IN REF. 2).
FT CONFLICT 1539 1539 F -> R (IN REF. 2).
FT CONFLICT 1546 1546 Q -> N (IN REF. 2).
SQ SEQUENCE 2133 AA; 239304 MW; 152BBA8997F570DA CRC64;
Query Match 82.7%; Score 720; DB 1; Length 2133;
Best Local Similarity 79.9%; Pred. No. 1.5e-61;
Matches 131; Conservative 20; Mismatches 13; Indels 0; Gaps 0;
QY 1 CDLNSCMPGLMESKATSDAQITASSYFTNMTWATNPSKARLHLQGRSNVNRQVNPKE 60
Db 1970 CDLNSCMPGLMGNKALSDSQITASSHLNIFATWPSQARLHLQGRTNWPRVSSAE 2029
QY 61 WLQVDFQKTMKVTGVTITQGVKSLTSMYKKEFLISSQDQHWTLFFQNGKVKVFGNQD 120
Db 2030 WLQVDLQKTVKVTGVTITQGVKSLTSMYKKEFLVSSQDGRWTLFLQDGHKVKVFGNQD 2089
QY 121 SFTFVNNSLDPPLLTLYRLRHPOSWVHQIALRMVGLGCEAQDLY 164
Db 2090 SFTFVNNSLDPPLLTLYRLRHPOSWVHQIALRMVGLGCEAQDLY 2133
RESULT 4
FAS_PIG STANDARD; PRT; 2258 AA.
AC Q9GLEP1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Coagulation factor V precursor (Activated protein C cofactor).
GN F5.
OS Sus scrofa (pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A., AND 3D-STRUCTURE MODELING OF F5/8 TYPE A AND C
 RP DOMAINS.
 RC TISSUE=Liver;
 RX MEDLINE=21121490; PubMed=11229814;
 RA Grimm D.R., Colter M.B., Braunschweig M., Alexander L.J., Neame P.J.,
 RA Kim H.K.W.;
 RT "Porcine factor V: cdna cloning, gene mapping, three-dimensional
 RT protein modeling of membrane binding sites and comparative anatomy of
 RT domains";
 RL Cell. Mol. Life Sci. 58:148-159(2001).
 CC -!- FUNCTION: Coagulation factor V is a cofactor that participates
 CC with factor Xa to activate prothrombin to thrombin.
 CC -!- SUBUNIT: Factor Va is composed of a heavy chain and a light
 CC chain, noncovalently bound. The interaction between the two chains
 CC is calcium-dependent.
 CC -!- DOMAIN: Domain B contains 41 X 9 AA tandem repeats. Domains C1
 CC and C2 may be involved in membrane binding.
 CC -!- PTM: Thrombin activates factor V proteolytically to the active
 CC cofactor, factor Va (formation of a heavy chain at the N-
 CC terminus and a light chain at the C-terminus).
 CC -!- SIMILARITY: Contains 3 F5/8 type A domains.
 CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
 CC -!- SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF191308; AAC28381.1; --
 DR HSPF; P12259; 1CZT.
 DR InterPro; IPR001117; Cu-oxidase.
 DR InterPro; IPR008972; Cupredoxin.
 DR InterPro; IPR00421; FAS8_C.
 DR InterPro; IPR008979; Gal_bind_like.
 DR Pfam; PF003194; Cu-oxidase; 3.
 DR Pfam; PF00754; F5 F8 type_C; 2.
 DR SMART; SM00231; FAS8C; 2.
 DR PROSITE; PS01285; FAS8C_1; 2.
 DR PROSITE; PS01286; FAS8C_2; 2.
 DR PROSITE; PS00022; FAS8C_3; 2.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
 KW Blood coagulation; Glycoprotein; Sulfation; Calcium; Signal; Zymogen;
 KW Repeat.
 FT SIGNAL. 1 22 POTENTIAL.
 FT CHAIN. 23 2258 COAGULATION FACTOR V.
 FT CHAIN. 23 737 COAGULATION FACTOR V HEAVY CHAIN (BY
 FT CHAIN. 23 737 SIMILARITY).
 FT PROPEP. 738 1611 ACTIVATION PEPTIDE (CONNECTING REGION)
 FT CHAIN. 1612 2258 (BY SIMILARITY).
 FT CHAIN. 1612 2258 COAGULATION FACTOR V LIGHT CHAIN (BY
 FT CHAIN. 1612 2258 SIMILARITY).
 FT DOMAIN. 30 329 F5/8 TYPE A 1.
 FT DOMAIN. 30 193 PLASTOCYANIN-LIKE 1.
 FT DOMAIN. 203 329 PLASTOCYANIN-LIKE 2.
 FT DOMAIN. 348 683 F5/8 TYPE A 2.
 FT DOMAIN. 348 525 PLASTOCYANIN-LIKE 3.
 FT DOMAIN. 535 683 PLASTOCYANIN-LIKE 4.
 FT DOMAIN. 691 1611 B.
 FT DOMAIN. 1168 1539 41 X 9 AA APPROXIMATE TANDEM REPEATS OF
 FT REPEAT. 1168 1176 T-L-S-P-D-L-[GS]-[HQ]-T.
 FT REPEAT. 1177 1185 1.
 FT REPEAT. 1186 1194 2.
 FT REPEAT. 1195 1203 3.
 FT REPEAT. 1195 1203 4.
 FT REPEAT. 1204 1212 5.
 FT
 FT 1221 REPEAT
 FT 1222 REPEAT
 FT 1230 REPEAT
 FT 1231 REPEAT
 FT 1239 REPEAT
 FT 1240 REPEAT
 FT 1248 REPEAT
 FT 1257 REPEAT
 FT 1258 REPEAT
 FT 1266 REPEAT
 FT 1275 REPEAT
 FT 1284 REPEAT
 FT 1285 REPEAT
 FT 1293 REPEAT
 FT 1302 REPEAT
 FT 1311 REPEAT
 FT 1312 REPEAT
 FT 1320 REPEAT
 FT 1321 REPEAT
 FT 1329 REPEAT
 FT 1330 REPEAT
 FT 1338 REPEAT
 FT 1347 REPEAT
 FT 1348 REPEAT
 FT 1356 REPEAT
 FT 1357 REPEAT
 FT 1365 REPEAT
 FT 1374 REPEAT
 FT 1383 REPEAT
 FT 1384 REPEAT
 FT 1392 REPEAT
 FT 1401 REPEAT
 FT 1410 REPEAT
 FT 1411 REPEAT
 FT 1419 REPEAT
 FT 1420 REPEAT
 FT 1428 REPEAT
 FT 1437 REPEAT
 FT 1446 REPEAT
 FT 1455 REPEAT
 FT 1464 REPEAT
 FT 1473 REPEAT
 FT 1482 REPEAT
 FT 1491 REPEAT
 FT 1492 REPEAT
 FT 1500 REPEAT
 FT 1501 REPEAT
 FT 1509 REPEAT
 FT 1510 REPEAT
 FT 1518 REPEAT
 FT 1527 REPEAT
 FT 1539 REPEAT
 FT 1541 REPEAT
 FT 1542 REPEAT
 FT 1543 REPEAT
 FT 1544 REPEAT
 FT 1545 REPEAT
 FT 1546 REPEAT
 FT 1547 REPEAT
 FT 1548 REPEAT
 FT 1549 REPEAT
 FT 1550 REPEAT
 FT 1551 REPEAT
 FT 1552 REPEAT
 FT 1553 REPEAT
 FT 1554 REPEAT
 FT 1555 REPEAT
 FT 1556 REPEAT
 FT 1557 REPEAT
 FT 1558 REPEAT
 FT 1559 REPEAT
 FT 1560 REPEAT
 FT 1561 REPEAT
 FT 1562 REPEAT
 FT 1563 REPEAT
 FT 1564 REPEAT
 FT 1565 REPEAT
 FT 1566 REPEAT
 FT 1567 REPEAT
 FT 1568 REPEAT
 FT 1569 REPEAT
 FT 1570 REPEAT
 FT 1571 REPEAT
 FT 1572 REPEAT
 FT 1573 REPEAT
 FT 1574 REPEAT
 FT 1575 REPEAT
 FT 1576 REPEAT
 FT 1577 REPEAT
 FT 1578 REPEAT
 FT 1579 REPEAT
 FT 1580 REPEAT
 FT 1581 REPEAT
 FT 1582 REPEAT
 FT 1583 REPEAT
 FT 1584 REPEAT
 FT 1585 REPEAT
 FT 1586 REPEAT
 FT 1587 REPEAT
 FT 1588 REPEAT
 FT 1589 REPEAT
 FT 1590 REPEAT
 FT 1591 REPEAT
 FT 1592 REPEAT
 FT 1593 REPEAT
 FT 1594 REPEAT
 FT 1595 REPEAT
 FT 1596 REPEAT
 FT 1597 REPEAT
 FT 1598 REPEAT
 FT 1599 REPEAT
 FT 1600 REPEAT
 FT 1601 REPEAT
 FT 1602 REPEAT
 FT 1603 REPEAT
 FT 1604 REPEAT
 FT 1605 REPEAT
 FT 1606 REPEAT
 FT 1607 REPEAT
 FT 1608 REPEAT
 FT 1609 REPEAT
 FT 1610 REPEAT
 FT 1611 REPEAT
 FT 1612 REPEAT
 FT 1613 REPEAT
 FT 1614 REPEAT
 FT 1615 REPEAT
 FT 1616 REPEAT
 FT 1617 REPEAT
 FT 1618 REPEAT
 FT 1619 REPEAT
 FT 1620 REPEAT
 FT 1621 REPEAT
 FT 1622 REPEAT
 FT 1623 REPEAT
 FT 1624 REPEAT
 FT 1625 REPEAT
 FT 1626 REPEAT
 FT 1627 REPEAT
 FT 1628 REPEAT
 FT 1629 REPEAT
 FT 1630 REPEAT
 FT 1631 REPEAT
 FT 1632 REPEAT
 FT 1633 REPEAT
 FT 1634 REPEAT
 FT 1635 REPEAT
 FT 1636 REPEAT
 FT 1637 REPEAT
 FT 1638 REPEAT
 FT 1639 REPEAT
 FT 1640 REPEAT
 FT 1641 REPEAT
 FT 1642 REPEAT
 FT 1643 REPEAT
 FT 1644 REPEAT
 FT 1645 REPEAT
 FT 1646 REPEAT
 FT 1647 REPEAT
 FT 1648 REPEAT
 FT 1649 REPEAT
 FT 1650 REPEAT
 FT 1651 REPEAT
 FT 1652 REPEAT
 FT 1653 REPEAT
 FT 1654 REPEAT
 FT 1655 REPEAT
 FT 1656 REPEAT
 FT 1657 REPEAT
 FT 1658 REPEAT
 FT 1659 REPEAT
 FT 1660 REPEAT
 FT 1661 REPEAT
 FT 1662 REPEAT
 FT 1663 REPEAT
 FT 1664 REPEAT
 FT 1665 REPEAT
 FT 1666 REPEAT
 FT 1667 REPEAT
 FT 1668 REPEAT
 FT 1669 REPEAT
 FT 1670 REPEAT
 FT 1671 REPEAT
 FT 1672 REPEAT
 FT 1673 REPEAT
 FT 1674 REPEAT
 FT 1675 REPEAT
 FT 1676 REPEAT
 FT 1677 REPEAT
 FT 1678 REPEAT
 FT 1679 REPEAT
 FT 1680 REPEAT
 FT 1681 REPEAT
 FT 1682 REPEAT
 FT 1683 REPEAT
 FT 1684 REPEAT
 FT 1685 REPEAT
 FT 1686 REPEAT
 FT 1687 REPEAT
 FT 1688 REPEAT
 FT 1689 REPEAT
 FT 1690 REPEAT
 FT 1691 REPEAT
 FT 1692 REPEAT
 FT 1693 REPEAT
 FT 1694 REPEAT
 FT 1695 REPEAT
 FT 1696 REPEAT
 FT 1697 REPEAT
 FT 1698 REPEAT
 FT 1699 REPEAT
 FT 1700 REPEAT
 FT 1701 REPEAT
 FT 1702 REPEAT
 FT 1703 REPEAT
 FT 1704 REPEAT
 FT 1705 REPEAT
 FT 1706 REPEAT
 FT 1707 REPEAT
 FT 1708 REPEAT
 FT 1709 REPEAT
 FT 1710 REPEAT
 FT 1711 REPEAT
 FT 1712 REPEAT
 FT 1713 REPEAT
 FT 1714 REPEAT
 FT 1715 REPEAT
 FT 1716 REPEAT
 FT 1717 REPEAT
 FT 1718 REPEAT
 FT 1719 REPEAT
 FT 1720 REPEAT
 FT 1721 REPEAT
 FT 1722 REPEAT
 FT 1723 REPEAT
 FT 1724 REPEAT
 FT 1725 REPEAT
 FT 1726 REPEAT
 FT 1727 REPEAT
 FT 1728 REPEAT
 FT 1729 REPEAT
 FT 1730 REPEAT
 FT 1731 REPEAT
 FT 1732 REPEAT
 FT 1733 REPEAT
 FT 1734 REPEAT
 FT 1735 REPEAT
 FT 1736 REPEAT
 FT 1737 REPEAT
 FT 1738 REPEAT
 FT 1739 REPEAT
 FT 1740 REPEAT
 FT 1741 REPEAT
 FT 1742 REPEAT
 FT 1743 REPEAT
 FT 1744 REPEAT
 FT 1745 REPEAT
 FT 1746 REPEAT
 FT 1747 REPEAT
 FT 1748 REPEAT
 FT 1749 REPEAT
 FT 1750 REPEAT
 FT 1751 REPEAT
 FT 1752 REPEAT
 FT 1753 REPEAT
 FT 1754 REPEAT
 FT 1755 REPEAT
 FT 1756 REPEAT
 FT 1757 REPEAT
 FT 1758 REPEAT
 FT 1759 REPEAT
 FT 1760 REPEAT
 FT 1761 REPEAT
 FT 1762 REPEAT
 FT 1763 REPEAT
 FT 1764 REPEAT
 FT 1765 REPEAT
 FT 1766 REPEAT
 FT 1767 REPEAT
 FT 1768 REPEAT
 FT 1769 REPEAT
 FT 1770 REPEAT
 FT 1771 REPEAT
 FT 1772 REPEAT
 FT 1773 REPEAT
 FT 1774 REPEAT
 FT 1775 REPEAT
 FT 1776 REPEAT
 FT 1777 REPEAT
 FT 1778 REPEAT
 FT 1779 REPEAT
 FT 1780 REPEAT
 FT 1781 REPEAT
 FT 1782 REPEAT
 FT 1783 REPEAT
 FT 1784 REPEAT
 FT 1785 REPEAT
 FT 1786 REPEAT
 FT 1787 REPEAT
 FT 1788 REPEAT
 FT 1789 REPEAT
 FT 1790 REPEAT
 FT 1791 REPEAT
 FT 1792 REPEAT
 FT 1793 REPEAT
 FT 1794 REPEAT
 FT 1795 REPEAT
 FT 1796 REPEAT
 FT 1797 REPEAT
 FT 1798 REPEAT
 FT 1799 REPEAT
 FT 1800 REPEAT
 FT 1801 REPEAT
 FT 1802 REPEAT
 FT 1803 REPEAT
 FT 1804 REPEAT
 FT 1805 REPEAT
 FT 1806 REPEAT
 FT 1807 REPEAT
 FT 1808 REPEAT
 FT 1809 REPEAT
 FT 1810 REPEAT
 FT 1811 REPEAT
 FT 1812 REPEAT
 FT 1813 REPEAT
 FT 1814 REPEAT
 FT 1815 REPEAT
 FT 1816 REPEAT
 FT 1817 REPEAT
 FT 1818 REPEAT
 FT 1819 REPEAT
 FT 1820 REPEAT
 FT 1821 REPEAT
 FT 1822 REPEAT
 FT 1823 REPEAT
 FT 1824 REPEAT
 FT 1825 REPEAT
 FT 1826 REPEAT
 FT 1827 REPEAT
 FT 1828 REPEAT
 FT 1829 REPEAT
 FT 1830 REPEAT
 FT 1831 REPEAT
 FT 1832 REPEAT
 FT 1833 REPEAT
 FT 1834 REPEAT
 FT 1835 REPEAT
 FT 1836 REPEAT
 FT 1837 REPEAT
 FT 1838 REPEAT
 FT 1839 REPEAT
 FT 1840 REPEAT
 FT 1841 REPEAT
 FT 1842 REPEAT
 FT 1843 REPEAT
 FT 1844 REPEAT
 FT 1845 REPEAT
 FT 1846 REPEAT
 FT 1847 REPEAT
 FT 1848 REPEAT
 FT 1849 REPEAT
 FT 1850 REPEAT
 FT 1851 REPEAT
 FT 1852 REPEAT
 FT 1853 REPEAT
 FT 1854 REPEAT
 FT 1855 REPEAT
 FT 1856 REPEAT
 FT 1857 REPEAT
 FT 1858 REPEAT
 FT 1859 REPEAT
 FT 1860 REPEAT
 FT 1861 REPEAT
 FT 1862 REPEAT
 FT 1863 REPEAT
 FT 1864 REPEAT
 FT 1865 REPEAT
 FT 1866 REPEAT
 FT 1867 REPEAT
 FT 1868 REPEAT
 FT 1869 REPEAT
 FT 1870 REPEAT
 FT 1871 REPEAT
 FT 1872 REPEAT
 FT 1873 REPEAT
 FT 1874 REPEAT
 FT 1875 REPEAT
 FT 1876 REPEAT
 FT 1877 REPEAT
 FT 1878 REPEAT
 FT 1879 REPEAT
 FT 1880 REPEAT
 FT 1881 REPEAT
 FT 1882 REPEAT
 FT 1883 REPEAT
 FT 1884 REPEAT
 FT 1885 REPEAT
 FT 1886 REPEAT
 FT 1887 REPEAT
 FT 1888 REPEAT
 FT 1889 REPEAT
 FT 1890 REPEAT
 FT 1891 REPEAT
 FT 1892 REPEAT
 FT 1893 REPEAT
 FT 1894 REPEAT
 FT 1895 REPEAT
 FT 1896 REPEAT
 FT 1897 REPEAT
 FT 1898 REPEAT
 FT 1899 REPEAT
 FT 1900 REPEAT
 FT 1901 REPEAT
 FT 1902 REPEAT
 FT 1903 REPEAT
 FT 1904 REPEAT
 FT 1905 REPEAT
 FT 1906 REPEAT
 FT 1907 REPEAT
 FT 1908 REPEAT
 FT 1909 REPEAT
 FT 1910 REPEAT
 FT 1911 REPEAT
 FT 1912 REPEAT
 FT 1913 REPEAT
 FT 1914 REPEAT
 FT 1915 REPEAT
 FT 1916 REPEAT
 FT 1917 REPEAT
 FT 1918 REPEAT
 FT 1919 REPEAT
 FT 1920 REPEAT
 FT 1921 REPEAT
 FT 1922 REPEAT
 FT 1923 REPEAT
 FT 1924 REPEAT
 FT 1925 REPEAT
 FT 1926 REPEAT
 FT 1927 REPEAT
 FT 1928 REPEAT
 FT 1929 REPEAT
 FT 1930 REPEAT
 FT 1931 REPEAT
 FT 1932 REPEAT
 FT 1933 REPEAT
 FT 1934 REPEAT
 FT 1935 REPEAT
 FT 1936 REPEAT
 FT 1937 REPEAT
 FT 1938 REPEAT
 FT 1939 REPEAT
 FT 1940 REPEAT
 FT 1941 REPEAT
 FT 1942 REPEAT
 FT 1943 REPEAT
 FT 1944 REPEAT
 FT 1945 REPEAT
 FT 1946 REPEAT
 FT 1947 REPEAT
 FT 1948 REPEAT
 FT 1949 REPEAT
 FT 1950 REPEAT
 FT 1951 REPEAT
 FT 1952 REPEAT
 FT 1953 REPEAT
 FT 1954 REPEAT
 FT 1955 REPEAT
 FT 1956 REPEAT
 FT 1957 REPEAT
 FT 1958 REPEAT
 FT 1959 REPEAT
 FT 1960 REPEAT
 FT 1961 REPEAT
 FT 1962 REPEAT
 FT 1963 REPEAT
 FT 1964 REPEAT
 FT 1965 REPEAT
 FT 1966 REPEAT
 FT 1967 REPEAT
 FT 1968 REPEAT
 FT 1969 REPEAT
 FT 1970 REPEAT
 FT 1971 REPEAT
 FT 1972 REPEAT
 FT 1973 REPEAT
 FT 1974 REPEAT
 FT 1975 REPEAT
 FT 1976 REPEAT
 FT 1977 REPEAT
 FT 1978 REPEAT
 FT 1979 REPEAT
 FT 1980 REPEAT
 FT 1981 REPEAT
 FT 1982 REPEAT
 FT 1983 REPEAT
 FT 1984 REPEAT
 FT 1985 REPEAT
 FT 1986 REPEAT
 FT 1987 REPEAT
 FT 1988 REPEAT
 FT 1989 REPEAT
 FT 1990 REPEAT
 FT 1991 REPEAT
 FT 1992 REPEAT
 FT 1993 REPEAT
 FT 1994 REPEAT
 FT 1995 REPEAT
 FT 1996 REPEAT
 FT 1997 REPEAT
 FT 1998 REPEAT
 FT 1999 REPEAT
 FT 2000 REPEAT
 FT 2001 REPEAT
 FT 2002 REPEAT
 FT 2003 REPEAT
 FT 2004 REPEAT
 FT 2005 REPEAT
 FT 2006 REPEAT
 FT 2007 REPEAT
 FT 2008 REPEAT
 FT 2009 REPEAT
 FT 2010 REPEAT
 FT 2011 REPEAT
 FT 2012 REPEAT
 FT 2013 REPEAT
 FT 2014 REPEAT
 FT 2015 REPEAT
 FT 2016 REPEAT
 FT 2017 REPEAT
 FT 2018 REPEAT
 FT 2019 REPEAT
 FT 2020 REPEAT
 FT 2021 REPEAT
 FT 2022 REPEAT
 FT 2023 REPEAT
 FT 2024 REPEAT
 FT 2025 REPEAT
 FT 2026 REPEAT
 FT 2027 REPEAT
 FT 2028 REPEAT
 FT 2029 REPEAT
 FT 2030 REPEAT
 FT 2031 REPEAT
 FT 2032 REPEAT
 FT 2033 REPEAT
 FT 2034 REPEAT
 FT 2035 REPEAT
 FT 2036 REPEAT
 FT 2037 REPEAT
 FT 2038 REPEAT
 FT 2039 REPEAT
 FT 2040 REPEAT
 FT 2041 REPEAT
 FT 2042 REPEAT
 FT 2043 REPEAT
 FT 2044 REPEAT
 FT 2045 REPEAT
 FT 2046 REPEAT
 FT 2047 REPEAT
 FT 2048 REPEAT
 FT 2049 REPEAT
 FT 2050 REPEAT
 FT 2051 REPEAT
 FT 2052 REPEAT
 FT 2053 REPEAT
 FT 2054 REPEAT
 FT 2055 REPEAT
 FT 2056 REPEAT
 FT 2057 REPEAT
 FT 2058 REPEAT
 FT 2059 REPEAT
 FT 2060 REPEAT
 FT 2061 REPEAT
 FT 2062 REPEAT
 FT 2063 REPEAT
 FT 2064 REPEAT
 FT 2065 REPEAT
 FT 2066 REPEAT
 FT 2067 REPEAT
 FT 2068 REPEAT
 FT 2069 REPEAT
 FT 2070 REPEAT
 FT 2071 REPEAT
 FT 2072 REPEAT
 FT 2073 REPEAT
 FT 2074 REPEAT
 FT 2075 REPEAT
 FT 2076 REPEAT
 FT 2077 REPEAT
 FT 2078 REPEAT
 FT 2079 REPEAT
 FT 2080 REPEAT
 FT 2081 REPEAT
 FT 2082 REPEAT
 FT 2083 REPEAT
 FT 2084 REPEAT
 FT 2085 REPEAT
 FT 2086 REPEAT
 FT 2087 REPEAT
 FT 2088 REPEAT
 FT 2089 REPEAT
 FT 2090 REPEAT
 FT 2091 REPEAT
 FT 2092 REPEAT
 FT 2093 REPEAT
 FT 2094 REPEAT
 FT 2095 REPEAT
 FT 2096 REPEAT
 FT 2097 REPEAT
 FT 2098 REPEAT
 FT 2099 REPEAT
 FT 2100 REPEAT
 FT 2101 REPEAT
 FT 2102 REPEAT
 FT 2103 REPEAT
 FT 2104 REPEAT
 FT 2105 REPEAT
 FT 2106 REPEAT
 FT 2107 REPEAT
 FT 2108 REPEAT
 FT 2109 REPEAT
 FT 2110 REPEAT
 FT 2111 REPEAT
 FT 2112 REPEAT
 FT 2113 REPEAT
 FT 2114 REPEAT
 FT 2115 REPEAT
 FT 2116 REPEAT
 FT 2117 REPEAT
 FT 2118 REPEAT
 FT 2119 REPEAT
 FT 2120 REPEAT
 FT 2121 REPEAT
 FT 2122 REPEAT
 FT 2123 REPEAT
 FT 2124 REPEAT
 FT 2125 REPEAT
 FT 2126 REPEAT
 FT 2127 REPEAT
 FT 2128 REPEAT
 FT 2129 REPEAT
 FT 2130 REPEAT
 FT 2131 REPEAT
 FT 2132 REPEAT
 FT 2133 REPEAT
 FT 2134 REPEAT
 FT 2135 REPEAT
 FT 2136 REPEAT
 FT 2137 REPEAT
 FT 2138 REPEAT
 FT 2139 REPEAT
 FT 2140 REPEAT
 FT 2141 REPEAT
 FT 2142 REPEAT
 FT 2143 REPEAT
 FT 2144 REPEAT
 FT 2145 REPEAT
 FT 2146 REPEAT
 FT 2147 REPEAT
 FT 2148 REPEAT
 FT 2149 REPEAT
 FT 2150 REPEAT
 FT 2151 REPEAT
 FT 2152 REPEAT
 FT 2153 REPEAT
 FT 2154 REPEAT
 FT 2155 REPEAT
 FT 2156 REPEAT
 FT 2157 REPEAT
 FT 2158 REPEAT
 FT 2159 REPEAT
 FT 2160 REPEAT
 FT 2161 REPEAT
 FT 2162 REPEAT
 FT 2163 REPEAT
 FT 2164 REPEAT
 FT 2165 REPEAT
 FT 2166 REPEAT
 FT 2167 REPEAT
 FT 2168 REPEAT
 FT 2169 REPEAT
 FT 2170 REPEAT
 FT 2171 REPEAT
 FT 2172 REPEAT
 FT 2173 REPEAT
 FT 2174 REPEAT
 FT 2175 REPEAT
 FT 2176 REPEAT
 FT 2177 REPEAT
 FT 2178 REPEAT
 FT 2179 REPEAT
 FT 2180 REPEAT
 FT 2181 REPEAT
 FT 2182 REPEAT
 FT 2183 REPEAT
 FT 2184 REPEAT
 FT 2185 REPEAT
 FT 2186 REPEAT
 FT 2187 REPEAT
 FT 2188 REPEAT
 FT 2189 REPEAT
 FT 2190 REPEAT
 FT 2191 REPEAT
 FT 2192 REPEAT
 FT 2193 REPEAT
 FT 2194 REPEAT
 FT 2195 REPEAT
 FT 2196 REPEAT
 FT 2197 REPEAT
 FT 2198 REPEAT
 FT 2199 REPEAT
 FT 2200 REPEAT
 FT 2201 REPEAT
 FT 2202 REPEAT
 FT 2203 REPEAT
 FT 2204 REPEAT
 FT 2205 REPEAT
 FT 2206 REPEAT
 FT 2207 REPEAT
 FT 2208 REPEAT
 FT 2209 REPEAT
 FT 2210 REPEAT
 FT 2211 REPEAT
 FT 2212 REPEAT
 FT 2213 REPEAT
 FT 2214 REPEAT
 FT 2215 REPEAT
 FT 2216 REPEAT
 FT 2217 REPEAT
 FT 2218 REPEAT
 FT 2219 REPEAT
 FT 2220 REPEAT
 FT 2221 REPEAT
 FT 2222 REPEAT
 FT 2223 REPEAT
 FT 2224 REPEAT
 FT 2225 REPEAT
 FT 2226 REPEAT
 FT 2227 REPEAT
 FT 2228 REPEAT
 FT 2229 REPEAT
 FT 2230 REPEAT
 FT 2231 REPEAT
 FT 2232 REPEAT
 FT 2233 REPEAT
 FT 2234 REPEAT
 FT 2235 REPEAT
 FT 2236 REPEAT
 FT 2237 REPEAT
 FT 2238 REPEAT
 FT 2239 REPEAT
 FT 2240 REPEAT
 FT 2241 REPEAT
 FT 2242 REPEAT
 FT 2243 REPEAT
 FT 2244 REPEAT
 FT 2245 REPEAT
 FT 2246 REPEAT
 FT 2247 REPEAT
 FT 2248 REPEAT
 FT 2249 REPEAT
 FT 2250 REPEAT
 FT 2251 REPEAT
 FT 2252 REPEAT
 FT 2253 REPEAT
 FT 2254 REPEAT
 FT 2255 REPEAT
 FT 2256 REPEAT
 FT 2257 REPEAT
 FT 2258 REPEAT
 FT 2259 REPEAT
 FT 2260 REPEAT
 FT 2261 REPEAT
 FT 2262 REPEAT
 FT 2263 REPEAT
 FT 2264 REPEAT
 FT 2265 REPEAT
 FT 2266 REPEAT
 FT 2267 REPEAT
 FT 2268 REPEAT
 FT 2269 REPEAT
 FT 2270 REPEAT
 FT 2271 REPEAT
 FT 2272 REPEAT
 FT 2273 REPEAT
 FT 2274 REPEAT
 FT 2275 REPEAT
 FT 2276 REPEAT
 FT 2277 REPEAT
 FT 2278 REPEAT
 FT 2279 REPEAT
 FT 2280 REPEAT
 FT 2281 REPEAT
 FT 2282 REPEAT
 FT 2283 REPEAT
 FT 2284 REPEAT
 FT 2285 REPEAT
 FT 2286 REPEAT
 FT 2287 REPEAT
 FT 2288 REPEAT
 FT 2289 REPEAT
 FT 2290 REPEAT
 FT 2291 REPEAT
 FT 2292 REPEAT
 FT 2293 REPEAT
 FT 2294 REPEAT
 FT 2295 REPEAT
 FT 2296 REPEAT
 FT 2297 REPEAT
 FT 2298 REPEAT
 FT 2299 REPEAT
 FT 2300 REPEAT
 FT 2301 REPEAT
 FT 2302 REPEAT
 FT 2303 REPEAT
 FT 2304 REPEAT
 FT 2305 REPEAT
 FT 2306 REPEAT
 FT 2307 REPEAT
 FT 2308 REPEAT
 FT 2309 REPEAT
 FT 2310 REPEAT
 FT 2311 REPEAT
 FT 2312 REPEAT
 FT 2313 REPEAT
 FT 2314 REPEAT
 FT 2315 REPEAT
 FT 2316 REPEAT
 FT 2317 REPEAT
 FT 2318 REPEAT
 FT 2319 REPEAT
 FT 2320 REPEAT
 FT 2321 REPEAT
 FT 2322 REPEAT
 FT 2323 REPEAT
 FT 2324 REPEAT
 FT 2325 REPEAT
 FT 2326 REPEAT
 FT 2327 REPEAT
 FT 2328 REPEAT
 FT 2329 REPEAT
 FT 2330 REPEAT
 FT 2331 REPEAT
 FT 2332 REPEAT
 FT 2333 REPEAT
 FT 2334 REPEAT
 FT 2335 REPEAT
 FT 2336 REPEAT
 FT 2337 REPEAT
 FT 2338 REPEAT
 FT 2339 REPEAT
 FT 2340 REPEAT
 FT 2341 REPEAT
 FT 2342 REPEAT
 FT 2343 REPEAT
 FT 2344 REPEAT
 FT 2345 REPEAT
 FT 2346 REPEAT
 FT 2347 REPEAT
 FT 2348 REPEAT
 FT 2349 REPEAT
 FT 2350 REPEAT
 FT 2351 REPEAT
 FT 2352 REPEAT
 FT 2353 REPEAT
 FT 2354 REPEAT
 FT 2355 REPEAT
 FT 2356 REPEAT
 FT 2357 REPEAT
 FT 2358 REPEAT
 FT 2359 REPEAT
 FT 2360 REPEAT
 FT 2361 REPEAT
 FT 2362 REPEAT
 FT 2363 REPEAT
 FT 2364 REPEAT
 FT 2365 REPEAT
 FT 2366 REPEAT
 FT 2367 REPEAT
 FT 2368 REPEAT
 FT 2369 REPEAT
 FT 2370 REPEAT
 FT 2

```

FT CARBOHYD 741 741 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 760 760 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 782 782 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 899 899 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 960 960 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1048 1048 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1057 1057 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1066 1066 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1073 1073 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1089 1089 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1174 1174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1480 1480 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1537 1537 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1597 1597 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1737 1737 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1886 1886 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2044 2044 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2243 2243 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 2258 AA; 256078 MW; 9159B9E0076A2ACC CRC64;

Query Match 44.2%; Score 385; DB 1; Length 2258;
Best Local Similarity 44.3%; Pred. No. 5e-29;
Matches 74; Conservative 33; Mismatches 54; Indels 6; Gaps 3;

QY 1 CDLNSCMPLGMSKAIISDAQITASSYFTNMPAT-WSPSKARLHLOGSNWAPQVNNPK 59
Db 2095 CEVNGCSTPLGMEGNIKEQITASSPKKSGWGDYWEFFRRLNAQGRVNAQAKANN 2154
QY 60 EWLQVDFQKTMKVTGVTQGVKLLTSYKFEFLISSQDGHQWTLFFQNGKV--KVFOG 117
Db 2155 QWLQIDLKIKKITAITTCGKSLSESEWYKRYTIQSDRGVEMKSKSVNVDKIFEG 2214
QY 118 NQDSFTPVNSLDPLTLRLVLRHPQSVHQAIRMEVLGCEAODLY 164
Db 2215 NNNKGVKNFFNPPIIRIFIRIPKMNQSIARLELFGC---DIY 2258

RESULT 5
FA5_BOVIN
ID_FAS_BOVIN STANDARD; PRT; 2211 AA.
AC Q28107; Q28108;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Coagulation factor V precursor (Activated protein C cofactor).
GN F5.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92147638; PubMed=1737753;
RA Guinjo E.R., Emon C.T., Mann K.G., Macgillivray R.T.;
RT "The complete cDNA sequence of bovine coagulation factor V.";
RL J. Biol. Chem. 267:2971-2978(1992).
RN [2]
RP DISULFIDE BONDS.
RX MEDLINE=95034740; PubMed=7947716;
RA Xue J., Kalafatis M., Silvera J.R., Kung C., Mann K.G.;
RT "Determination of the disulfide bridges in factor Va heavy chain.";
RL Biochemistry 33:13109-13116(1994).
CC -1- FUNCTION: Coagulation factor V is a cofactor that participates
CC with factor Xa to activate prothrombin to thrombin.
CC -1- SUBUNIT: Factor Va is composed of a heavy chain and a light
CC chain, noncovalently bound. The interaction between the two chains
CC is calcium-dependent.
CC -1- DOMAIN: Domain B contains 29.5 X 9 AA tandem repeats, and 2 X 14
CC AA repeats.

```

```

CC -1- PTM: Thrombin activates factor V proteolytically to the active
CC cofactor, factor Va (formation of a heavy chain at the N-
CC terminus and a light chain at the C-terminus).
CC -1- PTM: Sulfation is required for efficient thrombin cleavage and
CC activation and for full procoagulant activity (By similarity).
CC -1- SIMILARITY: Contains 3 F5/8 type A domains.
CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC -----
CC EMBL; M81440; AAA30512.1; -
CC EMBL; M81441; AAA30513.1; -
CC PIR; A42580; KFBOS.
CC HSP; P12259; ICZT.
CC InterPro; IPR001117; Cu-oxidase.
CC InterPro; IPR008972; Cupredoxin.
CC InterPro; IPR000421; FAS8_C.
CC InterPro; IPR008979; Gal_bind_like.
CC Pfam; PF00394; Cu-oxidase; 3.
CC Pfam; PF00754; F5_F8_type_C; 2.
CC SMART; SM00231; FAS8C; 2.
CC PROSITE; PS01285; FAS8C_1; 2.
CC PROSITE; PS01286; FAS8C_2; 2.
CC PROSITE; PS00022; FAS8C_3; 2.
CC PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
KW Blood coagulation; Glycoprotein; Sulfation; Calcium; Signal; Zymogen;
KW Repeat.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 2211 COAGULATION FACTOR V.
FT CHAIN 29 741 COAGULATION FACTOR HEAVY CHAIN (BY
FT PROPEP 742 1564 SIMILARITY).
FT CHAIN 1565 2211 ACTIVATION PEPTIDE (CONNECTING REGION)
FT CHAIN 30 327 COAGULATION FACTOR LIGHT CHAIN (BY
FT DOMAIN 30 193 SIMILARITY).
FT DOMAIN 203 327 F5/8 TYPE A 1.
FT DOMAIN 348 686 PLASTOCYANIN-LIKE 1.
FT DOMAIN 348 686 PLASTOCYANIN-LIKE 2.
FT DOMAIN 525 686 F5/8 TYPE A 2.
FT DOMAIN 535 686 PLASTOCYANIN-LIKE 3.
FT DOMAIN 596 1564 PLASTOCYANIN-LIKE 4.
FT DOMAIN 1124 1151 B.
FT DOMAIN 1124 1137 2 X 14 AA TANDEM REPEATS.
FT REPEAT 1138 1151 1-1.
FT REPEAT 1188 1453 1-2.
FT REPEAT 1188 1196 30 X 9 AA APPROXIMATE TANDEM REPEATS OF
FT REPEAT 1196 1205 [AS]-L-S-P-D-[LP]-[GS]-Q-[TE].
FT REPEAT 1197 1205 2-1.
FT REPEAT 1206 1214 2-2.
FT REPEAT 1215 1223 2-3.
FT REPEAT 1224 1232 2-4.
FT REPEAT 1233 1241 2-5.
FT REPEAT 1242 1250 2-6.
FT REPEAT 1251 1259 2-7.
FT REPEAT 1260 1268 2-8.
FT REPEAT 1269 1277 2-9.
FT REPEAT 1278 1286 2-10.
FT REPEAT 1287 1295 2-11.
FT REPEAT 1296 1304 2-12.
FT REPEAT 1305 1313 2-13.
FT REPEAT 1314 1322 2-14.
FT REPEAT 1323 1331 2-15.
FT REPEAT 1332 1340 2-16.
FT REPEAT 1341 1349 2-17.
FT REPEAT 1350 1358 2-18.
FT REPEAT 1350 1358 2-19.

```


MEDLINE=20052169; PubMed=10586886;
 RA Macedo-Ribeiro S., Bode W., Huber R., Quinn-Allen M.A., Kim S.W.,
 RA Ortel T.L., Bourenkov G.P., Bartunik H.D., Stubbs M.T., Kane W.H.,
 RA Fuentes-Prior P.;
 RT "Crystal structures of the membrane-binding C2 domain of human
 RT coagulation factor V.";
 RL Nature 402:434-439 (1999).
 [9]
 RP VARIANT MET-1764.
 RX MEDLINE=95179146; PubMed=7874144;
 RA Bayston T.A., Ireland H., Olds R.J., Thein S.L., Lane D.A.;
 RT "A polymorphism in the human coagulation factor V gene.";
 RL Hum. Mol. Genet. 3:2085-2085 (1994).
 [10]
 RP VARIANT APCR GLN-534.
 RX MEDLINE=94217810; PubMed=8164741;
 RA Bertina R.M., Koelenan B.P.C., Koster T., Rosendaal F.R.,
 RA Dirven R.J., de Ronde H., van der Velden P.A., Reitsma P.H.;
 RT "Mutation in blood coagulation factor V associated with resistance to
 RT activated protein C.";
 RL Nature 369:64-67 (1994).
 [11]
 RP VARIANTS ILE-1285 AND ARG-1327.
 RX MEDLINE=96351768; PubMed=8713778;
 RA Lunghi B., Iacoviello L., Gemmati D., Dilasio M.G., Castoldi E.,
 RA Pinotti M., Castaman G., Redaelli R., Mariani G., Marchetti G.,
 RA Bernardi F.;
 RT "Detection of new polymorphic markers in the factor V gene:
 RT association with factor V levels in plasma.";
 RL Thromb. Haemost. 75:45-48 (1996).
 [12]
 RP VARIANT APCR GLY-334, AND VARIANT LYS-513.
 RX MEDLINE=98122763; PubMed=9454741;
 RA Chan W.P., Lee C.K., Kwong Y.L., Lam C.K., Liang R.;
 RT "A novel mutation of Arg306 of factor V gene in Hong Kong Chinese.";
 RL Blood 91:1135-1139 (1998).
 [13]
 RP VARIANT APCR THR-334.
 RX MEDLINE=98122764; PubMed=9454742;
 RA Williamson D., Brown K., Luddington R., Baglin C., Baglin T.;
 RT "Factor V Cambridge: a new mutation (Arg306-to-Thr) associated with
 RT resistance to activated protein C.";
 RL Blood 91:1140-1144 (1998).
 [14]
 RP VARIANTS HIS-107; THR-413; LYS-513; SER-809; THR-817; ARG-859;
 RP ARG-865; GLU-925; GLN-1146; ALA-1530; SER-1685; VAL-1749; MET-1764;
 RP ILS-1820 AND GLY-2222, AND VARIANT APCR GLN-534.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemes J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 RT of human genes.";
 RL Nat. Genet. 22:231-238 (1999).
 [15]
 RP ERRATUM.
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemes J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
 RA Lander E.S.;
 RL Nat. Genet. 23:373-373 (1999).
 [16]
 RP VARIANT APCR HIS-2102.
 RX MEDLINE=21847288; PubMed=11858490;
 RA Schrijver I., Houissa-Kastally R., Jones C.D., Garcia X.C.,
 RA Zehnder J.L.;
 RT "Novel factor V C2-domain mutation (R2074H) in two families with
 RT factor V deficiency and bleeding.";
 RL Thromb. Haemost. 87:294-299 (2002).
 CC -1- FUNCTION: Coagulation factor V is a cofactor that participates
 CC with factor Xa to activate prothrombin to thrombin.
 CC -1- SUBUNIT: Factor Va is composed of a heavy chain and a light

chain, noncovalently bound. The interaction between the two chains
 is calcium-dependent.
 CC -1- DOMAIN: Domain B contains 35 x 9 AA tandem repeats, and 2 x 17 AA
 CC repeats.
 CC -1- PTM: Thrombin activates factor V proteolytically to the active
 CC cofactor, factor Va (formation of a heavy chain at the N-
 CC terminus and a light chain at the C-terminus).
 CC -1- PTM: Sulfation is required for efficient thrombin cleavage and
 CC activation and for full procoagulant activity.
 CC -1- DISEASE: Defects in F5 are the cause of Owren parahemophilia
 CC [MIM:227400], an hemorrhagic diastasis.
 CC -1- DISEASE: Defects in F5 are the cause of resistance to activated
 CC protein C (APCR) [MIM:188055], a form of thrombophilia. The APCR
 CC mutation is found in about 5% of the population which suggest that
 CC a slight thrombotic tendency may confer some advantage in fetal
 CC implantation.
 CC -1- SIMILARITY: Contains 3 F5/8 type A domains.
 CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
 CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L32779; AAB59401.1; JOINED.
 CC EMBL; L32755; AAB59401.1; JOINED.
 CC EMBL; L32756; AAB59401.1; JOINED.
 CC EMBL; L32757; AAB59401.1; JOINED.
 CC EMBL; L32758; AAB59401.1; JOINED.
 CC EMBL; L32759; AAB59401.1; JOINED.
 CC EMBL; L32760; AAB59401.1; JOINED.
 CC EMBL; L32761; AAB59401.1; JOINED.
 CC EMBL; L32762; AAB59401.1; JOINED.
 CC EMBL; L32763; AAB59401.1; JOINED.
 CC EMBL; L32764; AAB59401.1; JOINED.
 CC EMBL; L32765; AAB59401.1; JOINED.
 CC EMBL; L32766; AAB59401.1; JOINED.
 CC EMBL; L32767; AAB59401.1; JOINED.
 CC EMBL; L32768; AAB59401.1; JOINED.
 CC EMBL; L32769; AAB59401.1; JOINED.
 CC EMBL; L32770; AAB59401.1; JOINED.
 CC EMBL; L32771; AAB59401.1; JOINED.
 CC EMBL; L32772; AAB59401.1; JOINED.
 CC EMBL; L32773; AAB59401.1; JOINED.
 CC EMBL; L32774; AAB59401.1; JOINED.
 CC EMBL; L32775; AAB59401.1; JOINED.
 CC EMBL; L32776; AAB59401.1; JOINED.
 CC EMBL; L32777; AAB59401.1; JOINED.
 CC EMBL; L32778; AAB59401.1; JOINED.
 CC EMBL; M16967; AAB52424.1; --
 CC EMBL; M14335; AAB59532.1; --
 CC PDB; 1CZS; 26-NOV-99.
 CC PDB; 1CZT; 28-NOV-99.
 CC PDB; 1CZV; 28-NOV-99.
 CC PDB; 1FV4; 17-JAN-01.
 CC Genew; HGNC:3542; F5.
 CC MIM; 227400; --
 CC MIM; 134400; --
 CC MIM; 188055; --
 CC -----
 CC Query March 43.2%; Score 376; DB 1; Length 2224;
 CC Best Local Similarity 41.9%; Pred. No. 3,7e-28;
 CC Matches 70; Conservative 38; Mismatches 53; Indels 6; Gaps 3;
 QY 1 CDLNSCSPMLQMESKATSDAQITASSVFTNMAT-WSPSKARLHLQGRSNAPVNNPK 59
 DB 2061 CEVNGCSTPLGMENKIKENKQITASSFKSKSWGDYNEPPARLNAGGRVNAQKANNK 2120
 QY 60 EWLQVDFQKTMKVTGVTTCQVKSLTSMYVKEFLISSQDGHQWTLFFONGKV--KVFG 117

Db 2121 QWIEIDLLKKKTAITQCKSLSEMVKSYTHYSEGVWVPRKSSWDKIFEG 2180
QY 118 NODSFTFVNSLDPPLRLRYLRIHQSWVHQAIRMEVLGSCRAQDLY 164
Db 2181 NTWTKGHVKNFFNPPIISFIRVPTKNTQSIITRLLEFGC---DIY 2224
RESULT 7
MFGM RAT STANDARD; PRT; 427 AA.
AC P70450;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Lactadherin precursor (Milk fat globule-EGF factor 8) (MFG-E8) (O-
DE acetyl G33 ganglioside synthase) (AGS) (MFGM).
GN MFG-E8 OR AGS.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=96374422; PubMed=8780713;
RA Ogura K., Nara K., Watanabe Y., Kohno K., Tai T., Sanai Y.;
RX "Cloning and expression of cDNA for O-acetylation of G33
RT ganglioside.";
RL Biochem. Biophys. Res. Commun. 225:932-938(1996).
CC CC -1- FUNCTION: May be involved in phospholipid binding. Seems to
CC participate in the O-acetylation of G33 ganglioside stalic acid.
CC -1- SUBCELLULAR LOCATION: Peripheral membrane protein.
CC -1- TISSUE SPECIFICITY: Spleen, lung, heart, brain and muscle.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Contains 2 P5/8 type C domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D84068; BAA12210.1; -;
CC PIR; JC4915; JC4915.
CC HSSP; P00740; 1EDM.
CC InterPro; IPR006209; EGF like.
CC InterPro; IPR000421; FA58_C.
CC InterPro; IPR008979; Gal_bind_like.
CC InterPro; IPR006210; IEGF.
CC Pfam; PF00008; EGF; 2.
CC SMART; SM00754; F5_F8_type_C; 2.
CC SMART; SM00181; EGF_2.
CC SMART; SM00231; FA58C; 2.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; 2.
CC PROSITE; PS00026; EGF_3; 2.
CC PROSITE; PS01285; FA58C_1; 2.
CC PROSITE; PS01286; FA58C_2; 2.
CC PROSITE; PS00022; FA58C_3; 2.
KW Signal; Glycoprotein; Repeat; EGF-like domain; Milk.
FT SIGNAL 1 22
FT CHAIN 23 427
FT LACTADHERIN.
FT DOMAIN 24 61
FT EGF-LIKE 1.
FT DOMAIN 64 108
FT EGF-LIKE 2.
FT DOMAIN 111 267
FT F5/8 TYPE C 1.
FT DOMAIN 272 427
FT F5/8 TYPE C 2.
FT BY SIMILARITY.
FT DISULFID 28 39
FT BY SIMILARITY.
FT DISULFID 33 49
FT BY SIMILARITY.
FT DISULFID 51 60
FT BY SIMILARITY.
FT DISULFID 68 79

FT DISULFID 73 96 BY SIMILARITY.
FT DISULFID 98 107 BY SIMILARITY.
FT DISULFID 111 267 BY SIMILARITY.
FT DISULFID 254 258 BY SIMILARITY.
FT DISULFID 272 427 BY SIMILARITY.
FT SITE 87 89 CELL ATTACHMENT SITE (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 427 AA; 47413 MW; EA8C8631F3EB6047 CRC64;
Query Match 42.2%; Score 367.5; DB 1; Length 427;
Best Local Similarity 45.3%; Pred. No. 3.5e-28;
Matches 73; Conservative 28; Mismatches 57; Indels 3; Gaps 2;
QY 1 CDLNSCSNPLGMSKAISSDAQTASSYFT--NMFA-TWSPSKARLHLQGRSNWNRPOVNN 57
Db 267 CELHGCSSEPLGKKNTPDSQITASSSYKTNWLRFAFGWYHLGLDQGNKINWTAQNS 326
QY 58 PREWLQVDFQTKMTGVTTQGVKSLLTSMYVKEFLSSSQDGHQWTLFFQNGKVKVFG 117
Db 327 AKEWLQVLDGTQKKVTGIITQGFQGHQYVASYKVAHSDGQVQWTVYEQGTSKVFG 386
QY 118 NQDSFTFVNSLDPPLRLRYLRIHQSWVHQAIRMEVLGSC 158
Db 387 NLDNNSHKKNIFKPFMAYRVVLPFLSWHNRIITRLLELGC 427
RESULT 8
MFGM MOUSE
ID MFGM_MOUSE STANDARD; PRT; 463 AA.
AC P21956; P97800;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Lactadherin precursor (Milk fat globule-EGF factor 8) (MFG-E8) (MFGM)
DE (Sperm surface protein SP47) (WP47).
GN MFG-E8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1] SEQUENCE FROM N.A., AND SEQUENCE OF 23-35.
RP TISSUE=Mammary Gland;
RC MEDLINE=91046008; PubMed=2122462;
RX Stubbs J.D., Lekutis C., Singer K.L., Bui A., Yuzuki D.,
RA Srinivasan U., Parry G.;
RT "cDNA cloning of a mouse mammary epithelial cell surface protein
RT reveals the existence of epidermal growth factor-like domains linked
RT to factor VIII-like sequences".
RL Proc. Natl. Acad. Sci. U.S.A. 87:8417-8421(1990).
[2]
RN SEQUENCE OF 23-463 FROM N.A.
RP TISSUE=Testis;
RC Enselin M.A.;
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: May be involved in phospholipid binding. Zona pellucida-
CC binding protein.
CC -1- SUBCELLULAR LOCATION: Peripheral membrane protein.
CC -1- TISSUE SPECIFICITY: Mammary epithelial cell surfaces and
CC spermatozoan.
CC -1- DEVELOPMENTAL STAGE: mRNA expression is detectable in mammary
CC tissue from nonpregnant animals and maximal in the lactating
CC gland.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Contains 2 P5/8 type C domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL; M38337; AAA39534.1; -.
CC EMBL; Y11684; CAA72380.1; -.
CC PIR; A36479; A36479.
CC HSSP; P00740; IEDM.
CC MGD; MGI:102768; Mfge8.
CC InterPro; IPR001438; EGF II.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR000421; FA58_C.
CC InterPro; IPR008979; Gal_Bind_like.
CC InterPro; IPR006210; IEGF.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00754; F5 F8 type C; 2.
CC PRINTS; PR00010; EGFBLD-.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00231; FA58C; 2.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; 2.
CC PROSITE; PS00026; EGF_3; 2.
CC PROSITE; PS01285; FA58C_1; 2.
CC PROSITE; PS01285; FA58C_2; 2.
CC PROSITE; PS00022; FA58C_3; 2.
CC Signal; Glycoprotein; Repeat; EGF-like domain; Milk.
KW SIGNAL 1 22
FT CHAIN 23 463
FT DOMAIN 24 61
FT DOMAIN 64 108
FT DOMAIN 148 303
FT DOMAIN 308 463
FT SITE 87 89
FT DISULFID 28 39
FT DISULFID 33 49
FT DISULFID 51 60
FT DISULFID 68 79
FT DISULFID 73 96
FT DISULFID 98 107
FT DISULFID 148 303
FT DISULFID 290 463
FT DISULFID 308 463
FT CARBOHYD 61 61
FT CARBOHYD 266 266
FT CARBOHYD 316 316
FT CARBOHYD 426 426
FT CONFLICT 30 30
FT CONFLICT 35 35
FT CONFLICT 110 147
FT CONFLICT 168 168
FT CONFLICT 196 196
FT CONFLICT 309 309
FT CONFLICT 395 395
SQ SEQUENCE 463 AA; 51465 MW; D78B6C6EFBBA724D CRC64;

Query Match
Best Local Similarity 42.9%; Pred. No. 8.7e-27; Length 463;
Matches 69; Conservative 32; Mismatches 57; Indels 3; Gaps 2;

QY 1 CDLNCSEMPGLGMSKSAISDAQTASGYFT--NMFA-TWSPSKARLHLQGRSNAMKPVQNN 57
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
303 CELHGLEPLGLKNTIPDSQMSASSYKTNLRAFQWYPHLGRLDNQGKINAWTAQNS 362
QY 58 PKEWLQVDQKTKMVTGVTTQGVKSLTSMYKFEIISQSDGHQWTLFFQNGKVKVFG 117
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
363 AKEWLQVDLGTQGVTTGIIQTGARDPGHIOYVESYKVAHSDGVQWTVIEGSSKVFQ 422
QY 118 NQDSFTPVNSLDPPLLTRYLRHPSQWVHQIALRMEVLGC 158
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
423 NLDNNSHKKNIFKPPMARYRVRLPVSWHNRITLRLLELGC 463

```

```

RESULT 9
NRP2_RAT
ID_NRP2_RAT STANDARD; PRT; 925 AA.
AC O35276;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neuropilin-2 precursor (Vascular endothelial cell growth factor 165
DE receptor 2).
GN NRP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=97433085; PubMed=9288754;
RA Kolodkin A.L., Levengood D.V., Rowe E.G., Tai Y.-T., Giger R.J.,
RA Ginty D.D.;
RT "Neuropilin is a semaphorin III receptor.";
RL Cell 90:753-762(1997);
CC -!- FUNCTION: High affinity receptor for semaphorins 3C, 3P, VEGF-165
CC and VEGF-145 isoforms of VEGF, and the PLGF-2 isoform of PGF.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Found in certain neuronal populations of the
CC CNS and in other nonneuronal tissues including mesenchymal tissue
CC lining in the ribs.
CC -!- SIMILARITY: Belongs to the neuropilin family.
CC -!- SIMILARITY: Contains 2 CUB domains.
CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC -!- SIMILARITY: Contains 1 MAM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```



```
FT DISULFID 208 230 BY SIMILARITY.
FT DISULFID 277 427 BY SIMILARITY.
FT DISULFID 434 592 BY SIMILARITY.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 833 833 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 834 834 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 925 AA; 103896 MW; 3BF62903F644851C CRC64;

Query Match 37.8%; Score 329; DB 1; Length 925;
Best Local Similarity 43.4%; Pred. No. 4.7e-24;
Matches 69; Conservative 27; Mismatches 57; Indels 6; Gaps 3;

QY 6 CSMPLGMSKAISSDAQTASSYFTN--WATWSPSKARLHLQGRNAPQNNPKWQLQVD 65
DB CNAPLGMSGRIANEQIASSTFSD--GRWTPQOSRLH--GDGNGWTPNVDNKEYLQVD 332
QY 66 FQTKMTKVTGVTQGVKSILT--SMYKVEFLISSQDGHQWTLFPQNGKVVFGNODSFT 123
DB LRFUTMTATATQGAISRETQKGVYVKSYLEVSTNGEDMWYVHGHKHKVQFQANNDAT 392
QY 124 PVNSLDPPLTRYLRHPQSWHQAIALRMEVLGCEAQQ 162
DB LVLNKLTPELLTRIRIRPQWHLGIALRLFLPSCRVTDD 431

RESULT 10
MFGM_PIG STANDARD; PRT; 409 AA.
AC P79385;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Lactadherin (Milk fat globule-EGF factor 8) (MFG-E8) (MFGM) (Sperm
DE surface protein sp47) (pp47).
GN MFGE8.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Ensslin M.A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May be involved in phospholipid binding. Zona pellucida-
CC binding protein.
CC -!- SUBCELLULAR LOCATION: Peripheral membrane protein.
CC -!- TISSUE SPECIFICITY: Mammary epithelial cell surfaces and
CC spermatzoan.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y11683; CAA72379.1; -.
DR PIR; T11743; T11743.
DR HSSP; P00740; 1EDM.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000421; FAS5_C.
DR InterPro; IPR008579; Gal_bind_like.
DR InterPro; IPR006210; IEGF.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00231; FAS5C; 2.
```

```
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS01285; FAS5C_1; 2.
DR PROSITE; PS01286; FAS5C_2; 2.
DR PROSITE; PS00022; FAS5C_3; 2.
KW Glycoprotein; Repeat; EGF-like domain.
FT DOMAIN 2 41 EGF-LIKE 1.
FT DOMAIN 44 88 EGF-LIKE 2.
FT DOMAIN 91 247 F5/8 TYPE C 1.
FT DOMAIN 252 409 F5/8 TYPE C 2.
FT SITE 67 69 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 6 17 BY SIMILARITY.
FT DISULFID 11 29 BY SIMILARITY.
FT DISULFID 31 40 BY SIMILARITY.
FT DISULFID 91 247 BY SIMILARITY.
FT DISULFID 234 238 BY SIMILARITY.
FT DISULFID 252 409 BY SIMILARITY.
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 409 AA; 45725 MW; BOC07AF80029927A CRC64;

Query Match 37.7%; Score 328.5; DB 1; Length 409;
Best Local Similarity 39.9%; Pred. No. 2e-24;
Matches 65; Conservative 31; Mismatches 62; Indels 5; Gaps 2;

QY 1 CDLNSCSMPLGMSKAISSDAQTASSYFTN--WATWSPSKARLHLQGRNAPQNN 57
DB CELSGCAEPLGLKNDTIPNKQITASSFYRTWGLSAFSPFYARLDNQGKFWATQNS 306
QY 58 PXEWLQVDFQKTMKVTGVTQGVKSILTSMYKVEFLISSQDGHQWTLFPQNGKV--RVF 115
DB AEWLQIDLGSQRVTGIIITQCARDFGHIQVAAVKVAYSDDGVGVTYRQDGALEGIKIF 366
QY 116 QNQDSFTPVNSLDPPLTRYLRHPQSWHQAIALRMEVLGC 158
DB PGNLDDNSHKMKMFTPLTFVRILPVAWHNRITLRLVELLGC 409

RESULT 11
MFGV_HUMAN STANDARD; PRT; 387 AA.
AC Q08431;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Lactadherin precursor (Milk fat globule-EGF factor 8) (MFG-E8) (HMEG)
DE (Breast epithelial antigen BA46) (MFGM) [Contains: Medin].
GN MFGE8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast, and Breast carcinoma;
RX MEDLINE=9613308; PubMed=8639264;
RA Couto J.R., Taylor M.R., Godwin S.G., Ceriani R.L., Peterson J.A.;
RT "Cloning and sequence analysis of human breast epithelial antigen
RT BA46 reveals an RGD cell adhesion sequence presented on an epidermal
RT growth factor-like domain".
RT DNA Cell Biol. 15:281-286(1996).
RN [2]
RP SEQUENCE OF 170-387 FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=91371351; PubMed=1909932;
RA Larocca D., Peterson J.A., Urrea R., Kuniyoshi J., Bistrain A.M.,
RA Ceriani R.L.;
RT "A Mr 46,000 human milk fat globule protein that is highly expressed
RT in human breast tumors contains factor VIII-like domains."
RL Cancer Res. 51:4994-4998(1991).
RN [3]
RP PARTIAL SEQUENCE, AND CHARACTERIZATION.
```

RC TISSUE=Milk;
RX MEDLINE=98194924; PubMed=9535276;
RA Giuffrida M.G., Cavaletto M., Giunta C., Conti A.,
RA Godovac-Zimmermann J.;
RT "Isolation and characterization of full and truncated forms of human
RT breast carcinoma protein BA46 from human milk fat globule membranes.";
RL J. Protein Chem. 17:143-148(1998).
RN [4]
RP SEQUENCE OF 268-317, AND IDENTIFICATION OF MEDIN.
RX MEDLINE=99342076; PubMed=10411933;
RA Haeggqvist B., Naeslund J., Sletten K., Westermark G.T., Mucchiano G.,
RA Tjernberg L.O., Nordstedt C., Engstroem U., Westermark P.;
RT "Medin: an integral fragment of aortic smooth muscle cell-produced
RT lactadherin forms the most common human amyloid.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:8669-8674(1999).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=97405885; PubMed=9260929;
RA Taylor M.R., Couto J.R., Scallan C.D., Ceriani R.L., Peterson J.A.;
RT "Lactadherin (formerly BA46), a membrane-associated glycoprotein
RT expressed in human milk and breast carcinomas, promotes Arg-Gly-Asp
RT (RGD)-dependent cell adhesion.";
RL DNA Cell Biol. 16:861-869(1997).
CC CC
CC -!- FUNCTION: May be involved in phospholipid binding. Binds
CC specifically to rotavirus and inhibits its replication.
CC -!- FUNCTION: Medin is the main constituent of aortic medial amyloid.
CC -!- SUBCELLULAR LOCATION: Peripheral membrane protein.
CC -!- TISSUE SPECIFICITY: Mammary epithelial cell surfaces and aortic
CC media. Overexpressed in several carcinomas.
CC -!- PTM: MEDIN HAS A RAGGED N-TERMINUS WITH MINOR SPECIES STARTING AT
CC AMINO ACID 264 AND 273.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U58516; AAC50549.1; -;
CC EMBL: S56151; AAB19771.1; -;
CC PIR: A47285; A47285.
CC HSSP: P08709; 1BP9.
CC Genew: HGNC:7036; MFGE8.
CC MIN: 502381; -;
CC GO: GO:0007155; P:cell adhesion; TAG.
CC InterPro: IPR001438; EGF_II.
CC InterPro: IPR006209; EGF_like.
CC InterPro: IPR000421; FAS8 C.
CC InterPro: IPR008979; GalBind_like.
CC InterPro: IPR064210; IEGF.
CC Pfam: PF00008; EGF_1.
CC Pfam: PF00754; F5_F8type_C; 2.
CC PRINTS: PR00010; EGFBLD.
CC SMART: SM00181; EGF; 1.
CC SMART: SM00231; FAS8C; 2.
CC PROSITE: PS00022; EGF_1; 1.
CC PROSITE: PS01165; EGF_2; 1.
CC PROSITE: PS50025; EGF_3; 1.
CC PROSITE: PS01285; FAS8C_1; 2.
CC PROSITE: PS01286; FAS8C_2; 2.
CC PROSITE: PS50022; FAS8C_3; 2.
CC KW Signal; Glycoprotein; Milk; Repeat; EGF-like domain; Amyloid.
CC SIGNAL 1 23
CC CHAIN 24 387 LACTADHERIN.
CC CHAIN 202 387 LACTADHERIN; SHORT FORM.
CC CHAIN 268 317 MEDIN.
CC DOMAIN 24 67 EGF-LIKE.
CC DOMAIN 70 225 F5/8 TYPE C 1.
CC DOMAIN 230 387 F5/8 TYPE C 2.

FT SITE 46 48 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 27 38 BY SIMILARITY.
FT DISULFID 32 55 BY SIMILARITY.
FT DISULFID 57 66 BY SIMILARITY.
FT DISULFID 70 225 BY SIMILARITY.
FT DISULFID 212 216 BY SIMILARITY.
FT DISULFID 230 387 BY SIMILARITY.
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 387 AA; 43123 MW; 25E6571DEC83782D CEC64;
Query Match 37.6%; Score 327.5; DB 1; Length 387;
Best Local Similarity 40.2%; Pred. No. 2.3e-24;
Matches 66; Conservative 35; Mismatches 56; Indels 7; Gaps 3;
QY 1 CDLNSCSPLGWEKASIDAOITASSYF---TNMFATWSPSKARLHLCGRSNAWRPQVN 56
Db 225 CELNGCANPLGLKNNSIPDKOITASSSYKTKWGLHLP-SWNPYSYARLDKQGNFNAWAGSY 283
QY 57 NPKEWLQVDFQTKMKVTGVTTQGVKSLTSMYKFEISSLSSQDGHQWTLFF--QNGKVKV 114
Db 284 GNDQWLQVLDLGSKEVTGIITQAGRNFGSVQFVASYKVAYSNDSANWTEYQDPRTGSKI 343
QY 115 FQGNQDSFTFVNSLDPLLTTRYLRHPOSVHQAIRMEVLGC 158
Db 344 PFGNWDNHSKKNLFFETPILARYVRIPLVAWHNRILRLLELGC 387
RESULT 12
NRP2 HUMAN STANDARD; PRT; 931 AA.
AC O60462; O14820; O14821;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neuropilin-2 precursor (Vascular endothelial cell growth factor 165
DE receptor 2).
GN NRP2 OR VEGF165R2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A0 AND A17).
RX MEDLINE=97470888; PubMed=9331348;
RA Chen H., Chedotal A., He Z.-G., Goodman C.S., Tessier-Lavigne M.;
RT "Neuropilin-2, a novel member of the neuropilin family, is a high
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RT III.";
RL Neuron 19:547-559(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM A22).
RC TISSUE=Breast;
RX MEDLINE=98188099; PubMed=9529250;
RA Soker S., Takashima S., Miao H.-Q., Neufeld G., Klagsbrun M.;
RT "Neuropilin-1 is expressed by endothelial and tumor cells as an
RT isoform-specific receptor for vascular endothelial growth factor.";
RL Cell 92:735-745(1998).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=10748121;
RA Gluzman-Poltorak Z., Cohen T., Herzog Y., Neufeld G.;
RT "Neuropilin-1 are receptors for the 165-amino acid
RT form of vascular endothelial growth factor (VEGF) and of placenta
RT growth factor-2, but only neuropilin-2 functions as a receptor for
RT the 145-amino acid form of VEGF.";
RL J. Biol. Chem. 275:18040-18045(2000).
CC -!- FUNCTION: High affinity receptor for semaphorins 3C, 3F, VEGF-165
CC and VEGF-145 isoforms of VEGF, and the PLGF-2 isoform of PGF.
CC -!- SUBUNIT: Neuropilin-2 probably forms a heteromeric complex with
CC neuropilin-1 in order to be a functional semaphorin 3C receptor.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

EMBL; AF022856; AAC53379.1; -
 EMBL; AF022854; AAC53377.1; -
 EMBL; AF022855; AAC53378.1; -
 EMBL; AF022857; AAC53380.1; -
 EMBL; AF022858; AAC53381.1; -
 EMBL; AF022861; AAC53382.1; -
 HSP; P12259; 1CZT.
 MGD; MGI:1100492; NRP2.
 InterPro; IPR000859; CUB.
 InterPro; IPR000421; FA58_C.
 InterPro; IPR008979; Gal_Eind_like.
 InterPro; IPR000998; MAM_domain.
 Pfam; PF00431; CUB; 2.
 Pfam; PF00754; F5_F8_type_C; 2.
 PRINTS; PR00020; MAMDOMAIN.
 SMART; SM00042; CUB; 2.
 SMART; SM00231; FA58C; 2.
 SMART; SM00137; MAM; 1.
 PROSITE; PS01180; CUB; 2.
 PROSITE; PS01285; FA58C_1; 2.
 PROSITE; PS01286; FA58C_2; 2.
 PROSITE; PS00022; FA58C_3; 2.
 PROSITE; PS00060; MAM_2; 1.
 Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
 Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 21 931
 FT DOMAIN 21 864
 FT TRANSMEM 865 889
 FT DOMAIN 890 931
 FT DOMAIN 28 142
 FT DOMAIN 149 267
 FT DOMAIN 277 427
 FT DOMAIN 434 592
 FT DOMAIN 642 802
 FT DOMAIN 838 845
 FT DISULFID 28 55
 FT DISULFID 83 105
 FT DISULFID 149 175
 FT DISULFID 208 230
 FT DISULFID 277 427
 FT DISULFID 434 592
 FT CARBOHYD 152 157
 FT CARBOHYD 157 157
 FT CARBOHYD 629 629
 FT CARBOHYD 839 839
 FT VARSPPLIC 809 813
 FT VARSPPLIC 809 830
 FT VARSPPLIC 814 830
 FT VARSPPLIC 810 931
 ZDFKESWLYTDLPIITIIAMSSGLVLLGACGLLYCT
 SSGLSRSCSTLTENFELYDGLKHKVINKQCCSEA
 -> GTLPFGTEPTVDFVQPIPAYWYVNAAGAVLVLAS
 VVALVLYHRFRVAAKTDHSTYKTYTNGAPLAVET
 LTKLEQERSHC (in isoform B0).
 /FTid=VSP_004346
 VDIPTFGGEGYDEIDDEYEGDWSNSSSTSGAGPSGK
 EKSWLYTDLPIITIIAMSSGLVLLGACGLLYCTCSYS
 GLSRSCSTLTENFELYDGLKHKVINKQCCSEA -> G
 GTLPFGTEPTVDFVQPIPAYWYVNAAGAVLVLASVVL

FT ALVLVLYHRFRVAAKTDHSTYKTYTNGAPLAVETPLTI
 FT KLEQERSHC (in isoform B5).
 FT /FTid=VSP_004347.
 FT G -> I (IN REF. 1; AAC53380/AAC53381).
 SQ SEQUENCE 931 AA; 104558 MW; 76F2443F411D2F63 CRC64;
 Query Match 37.5%; Score 327; DB 1; Length 931;
 Best Local Similarity 42.1%; Pred. No. 7.4e-24;
 Matches 67; Conservative 30; Mismatches 56; Indels 6; Gaps 3;
 QY 6 CSMPLGMSKATSDAQITASSVFTNMFATSPSKARLHLOGRSNAWRQVANNPKXEWLQVD 65
 DB 277 CNVPLGMSGRIANEQISASTFSD--GRWTQQSRLH--GDDNGWTENLDSNKKEYLQVD 332
 QY 66 FOKTKVTVGTQGVKSLT--SMVVKFLLSISSODGHWTLFFQNGKVKVQGNQDSFT 123
 DB 333 LFLFLMLTAIATQGAISRETQKGYVYSKYLEVTNGEDMMVYRHGKHKIFQANNDATE 392
 QY 124 PVVNSLDPLLTTRYLRIHPQSWHQAIALRMEVLGCEAOD 162
 DB 393 VVNLKMLPLLTFRIRIRPQTHLGLIALRLFLFGCRVTD 431
 RESULT 14
 MFCM BOVIN
 ID MFCM BOVIN STANDARD; PRT; 427 AA.
 AC Q95114; P79344; Q27959;
 DT 01-NOV-1997 (Rel. 35; Created)
 DT 01-NOV-1997 (Rel. 35; Last sequence update)
 DT 15-MAR-2004 (Rel. 43; Last annotation update)
 DE Lactadherin precursor (Milk fat globule-EGF factor 8) (MFG-E8)
 DE (MGP57/53) (PAS-6/PAS-7 glycoprotein) (MFGM) (Sperm surface protein
 DE SP47) (BP47) (Components 15/16).
 GN MFG8.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]_TaxID=9913;
 RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.
 RC STRAIN=Holstein; TISSUE=Mammary gland;
 RX MEDLINE=97008954; PubMed=8856064;
 RA Hvarregaard J., Andersen M.H., Berglund L., Rasmussen J.T.,
 RA Petersen T.E.;
 RT "Characterization of glycoprotein PAS-6/7 from membranes of bovine
 RT milk fat globules";
 RL Eur. J. Biochem. 240:628-636 (1996).
 RN [2]
 RP SEQUENCE OF 19-427 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=96125736; PubMed=8541316;
 RA Aoki N., Kishi M., Taniguchi Y., Adachi T., Nakamura R.,
 RA Matsuda T.;
 RT "Molecular cloning of glycoprotein antigens MGP57/53 recognized by
 RT monoclonal antibodies raised against bovine milk fat globule
 RT membrane";
 RL Biochim. Biophys. Acta 1245:385-391 (1995).
 RN [3]
 RP SEQUENCE OF 19-427 FROM N.A.
 RC TISSUE=Testis;
 RA Ensalin M.A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 140-146; 174-187; 233-246 AND 422-427.
 RC TISSUE=Milk;
 RX MEDLINE=93250576; PubMed=8485470;
 RA Mather I.H., Banghart L.R., Lane W.S.;
 RT "The major fat-globule membrane proteins, bovine components 15/16 and
 RT guinea-pig GP 55, are homologous to MGP-E8, a murine glycoprotein
 RT containing epidermal growth factor-like and factor V/VIII-like
 RT sequences";
 RL Biochem. Mol. Biol. Int. 29:545-554 (1993).

CC -!- FUNCTION: Probably associates with phospholipids on the surface of
 CC mammary epithelial cells and milk fat globules. Zona pellucida-
 CC binding protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC isoId=Q95114-1; Sequence=Displayed;
 CC Name=Short;
 CC isoId=Q95114-2; Sequence=VSP_001398;
 CC -!- TISSUE SPECIFICITY: Milk and spermatozoan.
 CC -!- PTM: THE 2 O-LINKED GLYCANS CONSTITUTE OF GAL, GLCNAC AND FUC, WITH
 CC PROBABLY FUC AS REDUCING TERMINAL SUGAR.
 CC -!- SIMILARITY: Contains 2 EGF-like domains.
 CC -!- SIMILARITY: Contains 2 EGF-like domains.
 CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X91895; CAA62997.1; -
 CC EMBL; S80643; AAB35894.2; -
 CC EMBL; Y11719; CAA72406.1; -
 CC PIR; S74211; S74211.
 CC HSP; P00740; I1XA.
 CC InterPro; IPR006209; EGF like.
 CC InterPro; IPR000421; FAS8_C.
 CC InterPro; IPR008979; Gal_Bind_like.
 CC InterPro; IPR006210; IEGF.
 CC Pfam; PF00008; EGF; 2.
 CC Pfam; PF00754; F5_F8_type_C; 2.
 CC SMART; SM00181; EGF; 2.
 CC SMART; SM00231; FAS8C; 2.
 CC PROSITE; PS00022; EGF_1; 2.
 CC PROSITE; PS01186; EGF_2; 2.
 CC PROSITE; PS00026; EGF_3; 2.
 CC PROSITE; PS01285; FAS8C_1; 2.
 CC PROSITE; PS01285; FAS8C_2; 2.
 CC PROSITE; PS01286; FAS8C_3; 2.
 CC PROSITE; PS00022; FAS8C_3; 2.
 CC Signal; Glycoprotein; Milk; Repeat; EGF-like domain;
 CC Alternative splicing.
 CC SIGNAL 1
 CC CHAIN 19 427 LACTADHERIN.
 CC DOMAIN 20 59 EGF-LIKE 1.
 CC DOMAIN 62 106 EGF-LIKE 2.
 CC DOMAIN 109 265 F5/8 TYPE C 1.
 CC DOMAIN 270 427 F5/8 TYPE C 2.
 CC SITE 85 87 CELL ATTACHMENT SITE (POTENTIAL).
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC
 CC O-LINKED (FUC...) (IN PAS-6).
 CC O-LINKED (FUC...) (IN PAS-7).
 CC N-LINKED (GLCNAC...) (HYBRID)
 CC (IN PAS-6 AND PAS-7).
 CC N-LINKED (GLCNAC...) (HIGH MANNOSIDE)
 CC (IN PAS-6).
 CC Missing (in isoform Short).
 CC /FTid=VSP_001398.
 CC A -> F (IN REF. 1).
 CC L -> Q (IN REF. 1).
 CC CONFLICT 28 28
 CC SEQUENCE 427 AA; 47411 MW; 4CBEBE3A1DC4EB24 CRC64;
 CC
 CC Query Match 37.3%; Score 324.5; DB 1; Length 427;

Best Local Similarity 39.3%; Pred. No. 5.1e-24;
 Matches 64; Conservative 29; Mismatches 65; Indels 5; Gaps 2;
 QY 1 CDLNSCMPILGMSKAISSDAQTASSYFTN--MFATWSPSKARLHLCGRSNARWQVNN 57
 Db 265 CELNGCTEPLGLKNDTIPKQITASSYFTWGLSAPFWFFYARLDNQGRFNAQTNS 324
 QY 58 PREWLQVDFQKTMKVTGVTQGVKSLTSMYKFEFLISSQDGHOWTLFFQNG--KVKVF 115
 Db 325 ASEWLQIDLGSKQKRVGTGIIQTQARDFGHGYVAAVAVAYDGGVTWTEYKDPCASESKIP 384
 QY 116 QGNQDSFTPVNSLDPPLTLRLVLRHPQSWVHQIALMEVLGC 158
 Db 385 PGNVNDNSHKKNIFETPFQARFVRIOQVAVAHNRITLURVELLC 427
 RESULT 15
 NRPI_CHICK
 ID NRPI_CHICK STANDARD; PRT; 914 AA.
 AC P79735;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Neurotrophin-1 precursor (A5 protein).
 GN NRPI OR NRP.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=White Leghorn; TISSUE=Embryonic brain;
 RX MEDLINE=95324761; PubMed=7601310;
 RA Takagi S., Kasuya Y., Shimizu M., Matsuura T., Tsuboi M., Kawakami A.,
 RA Fujisawa H.;
 RT "Expression of a cell adhesion molecule, neuropilin, in the
 RT developing chick nervous system.";
 RL Dev. Biol. 170:207-222(1995).
 CC -!- FUNCTION: Receptor involved in the development of the
 CC cardiovascular system, in angiogenesis, in the formation of
 CC certain neuronal circuits and in organogenesis outside the nervous
 CC system. It mediates the chemorepulsant activity of semaphorins (By
 CC similarity). Seems to have calcium-independent cell adhesion
 CC properties.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Developing nervous system; optic tectum
 CC (layers D and E of SGFS), amacrine cells of retina, neurites of
 CC dorsal root ganglia. Also expressed in nonneuronal cells, e.g.
 CC blood vessels in the entire embryo.
 CC -!- SIMILARITY: Belongs to the neuropilin family.
 CC -!- SIMILARITY: Contains 2 CUB domains.
 CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
 CC -!- SIMILARITY: Contains 1 MAM domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D45416; BAA08256.1; -
 CC HSP; P12259; 1CZT.
 CC InterPro; IPR000859; CUB.
 CC InterPro; IPR000421; FAS8_C.
 CC InterPro; IPR008979; Gal_Bind_like.
 CC InterPro; IPR000998; MAM_domain.
 CC Pfam; PF00431; CUB; 2.
 CC Pfam; PF00754; F5_F8_type_C; 2.
 CC Pfam; PF00629; MAM; 1.
 CC PRINTS; PR00020; MAMDOMAIN.

DR SMART; SM00042; CUB; 2.
 DR SMART; SM00231; FA58C; 2.
 DR SMART; SM00137; MAM; 1.
 DR PROSITE; PS01180; CUB; 2.
 DR PROSITE; PS01285; FA58C_1; 2.
 DR PROSITE; PS01286; FA58C_2; 2.
 DR PROSITE; PS00022; FA58C_3; 2.
 DR PROSITE; PS00740; MAM_1; 1.
 DR PROSITE; PS00060; MAM_2; 1.
 KW Angiogenesis; Transmembrane; Glycoprotein; Neurone; Signal; Repeat;
 KW Receptor; Cell adhesion.
 FT SIGNAL 1 18
 FT CHAIN 19 914
 FT DOMAIN 20 847
 FT TRANSMEM 848 870
 FT DOMAIN 871 914
 FT DOMAIN 25 139
 FT DOMAIN 145 263
 FT DOMAIN 273 422
 FT DOMAIN 429 581
 FT DOMAIN 536 801
 FT DISULFID 25 52
 FT DISULFID 80 102
 FT DISULFID 145 171
 FT DISULFID 204 226
 FT DISULFID 273 422
 FT DISULFID 429 581
 SQ SEQUENCE 914 AA; 102480 MW; DD2EE6D6F0CBB68C CRC64;
 Query Match 31.6%; Score 275.5; DB 1; Length 914;
 Best Local Similarity 37.7%; Pred. No. 7.2e-19;
 Matches 60; Conservative 24; Mismatches 68; Indels 7; Gaps 3;
 QY 6 CSMPLGMSKASDAQITASSYFTNMFATWSPSKARLHLOGRSNAWRPQVNNPKWLQVD 65
 Db 273 CMEPLGMSGSIHQDITVSSQYS---AIWSSERSLNYP--ENGWTPGEDSVREWIQVD.327
 QY 66 FQKTMKVTGVTTQGV--KSLITSMYVKEFLISSODGHQWTLRFONGKVKVFGQNDST 123
 Db 328 LGLRFVSGIGTQGAISKETKVEYLYKTYRVDVSSNGEDWITLKEGNKPVVFGQNSNPTD 387
 QY 124 PVVNSLDPPLLRHLPQSWHQAIALRMEVLGCEAOD 162
 Db 388 VVYRPFKPKVLTFRVRIKPVSWENGVSRLRFVYGVCKITD 426

Search completed: September 29, 2004, 16:40:02
 Job time : 14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 16:34:18 ; Search time 41 Seconds
(without alignments)
1262.073 Million cell updates/sec

Title: US-10-049-399A-1_COPY_2169_2332

Perfect score: 871

Sequence: 1 CDLNSCMPGMSKALSDA.....VWHQIALRMEVLGCEAQQDIY 164

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	871	100.0	216	4 Q14286	Q14286 homo sapien
2	736	84.5	2343	6 O18806	O18806 canis fami
3	731	83.9	2343	6 O62730	O62730 canis fami
4	618	71.0	2258	11 Q7TN96	Q7TN96 rattus norv
5	568	65.2	1377	13 Q804X3	Q804X3 gallus gall
6	465	53.4	1639	13 Q804W6	Q804W6 fugu rubrip
7	405	46.5	1460	13 Q7SZN0	Q7SZN0 pseudonaja
8	402	46.2	745	13 Q804X4	Q804X4 gallus gall
9	390	44.8	2183	11 Q88783	Q88783 mus musculu
10	385	44.2	2102	11 Q7TPK2	Q7TPK2 rattus norv
11	365.5	42.1	480	4 Q8N610	Q8N610 homo sapien
12	366.5	42.1	480	4 O4385A	O4385A homo sapien
13	365.5	42.0	407	13 Q8AIE0	Q8AIE0 brachydanio
14	365.5	42.0	2119	13 Q90X47	Q90X47 brachydanio
15	357.5	41.0	426	11 Q9WTS3	Q9WTS3 mus musculu
16	357.5	41.0	463	11 Q9RIX9	Q9RIX9 mus musculu

17	357.5	41.0	470	11	Q8C4U8	Q8C4U8 mus musculu
18	357.5	41.0	480	11	Q8C9F7	Q8C9F7 mus musculu
19	357.5	41.0	480	11	O35474	O35474 mus musculu
20	354.5	40.7	1802	13	Q804W5	Q804W5 fugu rubrip
21	330	37.9	384	11	Q8C8K0	Q8C8K0 mus musculu
22	327	37.5	555	4	Q9H2E2	Q9H2E2 homo sapien
23	327	37.5	901	4	Q9H2E4	Q9H2E4 homo sapien
24	327	37.5	901	4	Q9H2D5	Q9H2D5 homo sapien
25	327	37.5	906	4	Q9H2D4	Q9H2D4 homo sapien
26	327	37.5	906	4	Q9H2E3	Q9H2E3 homo sapien
27	327	37.5	926	11	Q8QZT7	Q8QZT7 mus musculu
28	327	37.5	931	4	Q7Z3T9	Q7Z3T9 homo sapien
29	323.5	37.1	312	4	Q7Z3D2	Q7Z3D2 homo sapien
30	312	35.8	905	13	Q80014	Q80014 gallus gall
31	311	35.7	919	13	Q8UVR0	Q8UVR0 gallus gall
32	311	35.7	936	13	Q8UVQ9	Q8UVQ9 gallus gall
33	306.5	35.2	743	4	Q8N6M4	Q8N6M4 homo sapien
34	306.5	35.2	775	4	Q8TDX2	Q8TDX2 homo sapien
35	306.5	35.2	775	4	Q86PD2	Q86PD2 homo sapien
36	299.5	34.4	769	11	Q91ZV2	Q91ZV2 rattus norv
37	294.5	33.8	602	11	Q8BK14	Q8BK14 mus musculu
38	294.5	33.8	769	11	Q91ZV3	Q91ZV3 mus musculu
39	285.5	32.8	524	4	Q8N8Z6	Q8N8Z6 homo sapien
40	285.5	32.8	539	4	Q81YK5	Q81YK5 homo sapien
41	283.5	32.5	858	5	O76470	O76470 lytechinus
42	278.5	32.0	923	13	Q8QFX6	Q8QFX6 brachydanio
43	278.5	32.0	923	13	Q8AXP1	Q8AXP1 brachydanio
44	273	31.3	51	4	Q9UQQ5	Q9UQQ5 homo sapien
45	272	31.2	335	4	Q9BTL9	Q9BTL9 homo sapien

ALIGNMENTS

RESULT 1

Q14286 PRELIMINARY; PRT; 216 AA.
 ID Q14286 AC Q14286; DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DE 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Coagulation factor VIII associated protein B (Coagulation factor VIII, DE procoagulant component) (Hemophilia A).
 GN F8B.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93052386; PubMed=1427887;
 RA Levinson B., Kenwick S., Gamel P., Fisher K., Gitschler J.;
 RT "Evidence for a third transcript from the human factor VIII gene.";
 RL Genomics 14:585-589(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
 DE EMBL; M90707; AAA58466.1; -.
 DR EMBL; BC022513; AAH22513.1; -.
 DR PIR; A44259; A44258.
 DR HSSP; P00451; ICFC.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR000421; FAS8C.
 DR InterPro; IPR008979; GalBind-like.
 DR Pfam; PF00754; F5_F8 type C; 2.
 DR SMART; SM00231; FAS8C_1_1.
 DR PROSITE; PS01285; FAS8C_1; 1.
 DR PROSITE; PS01286; FAS8C_2; 2.
 DR PROSITE; PS50022; FAS8C_3; 2.
 SQ SEQUENCE 216 AA; 24641 MW; 6C82D4F89E35A376 CRC64;

```

Query Match
Best Local Similarity 100.0%; Score 871; DB 4; Length 216;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCMPLGMSKAISSDAQITASSYFTNMATWSPSKARLHLQGRSNARWPQVNNPKE 60
DB 53 CDLNSCMPLGMSKAISSDAQITASSYFTNMATWSPSKARLHLQGRSNARWPQVNNPKE 112

QY 61 WLQVDFOKTMKVTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGNQD 120
DB 113 WLQVDFOKTMKVTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGNQD 172

QY 121 SFTPVNSLDPLLTRYLRIHPQSVWHQIALRMEVLGCEAQDLY 164
DB 173 SFTPVNSLDPLLTRYLRIHPQSVWHQIALRMEVLGCEAQDLY 216

RESULT 2
OL8806 PRELIMINARY; PRT; 2343 AA.
AC OL8806;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Factor VIII.
GN F8.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Cameron C., Nottley C., Hoyle S., McGlynn L., Hough C., Kamisue S.,
RA Giles A., Lilliecrap D.;
RT "The canine factor VIII cDNA and 5' flanking sequence.";
RL Submitted (NCV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
DR EMBL; AF016234; AAB87412.1; -.
DR HSSP; P00451; 1CFG.
DR GO; GO:0005507; F: copper ion binding; IEA.
DR GO; GO:0007155; P: cell adhesion; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR000421; FA58_C.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00231; FA58C_2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS50022; FA58C_3; 2.
DR PROSITE; PS50079; MULTICOPPER OXIDASE1; 3.
DR PROSITE; PS50079; MULTICOPPER OXIDASE1; 3.
SQ SEQUENCE 2343 AA; 265829 MW; A854FAE571C3B399 CRC64;

Query Match
Best Local Similarity 84.5%; Score 736; DB 6; Length 2343;
Matches 134; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 CDLNSCMPLGMSKAISSDAQITASSYFTNMATWSPSKARLHLQGRSNARWPQVNNPKE 60
DB 2180 CDLNSCMPLGMSKAISSDAQITASSYFTNMATWSPSKARLHLQGRSNARWPQVNNPKE 2239

QY 61 WLQVDFOKTMKVTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGNQD 120
DB 2240 WLQVDFOKTMKVTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGNQD 2299

QY 121 SFTPVNSLDPLLTRYLRIHPQSVWHQIALRMEVLGCEAQ 161
DB 2300 SFTPVNSLDPLLTRYLRIHPQSVWHQIALRMEVLGCEAQ 2340

Query Match
Best Local Similarity 83.2%; Pred. No. 3.5e-66;
Matches 134; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

QY 1 CDLNSCMPLGMSKAISSDAQITASSYFTNMATWSPSKARLHLQGRSNARWPQVNNPKE 60
DB 2180 CDLNSCMPLGMSKAISSDAQITASSYFTNMATWSPSKARLHLQGRSNARWPQVNNPKE 2239

QY 61 WLQVDFOKTMKVTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGNQD 120
DB 2240 WLQVDFOKTMKVTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGNQD 2299

QY 121 SFTPVNSLDPLLTRYLRIHPQSVWHQIALRMEVLGCEAQ 161
DB 2300 SFTPVNSLDPLLTRYLRIHPQSVWHQIALRMEVLGCEAQ 2340

```

```

RESULT 3
O62730 PRELIMINARY; PRT; 2343 AA.
AC O62730;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Factor VIII.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney, and Spleen;
RA Gordy P.W., Bowen R.A.;
RT "Characterization of the canine factor VIII cDNA.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
DR EMBL; AF049489; AAC05384.1; -.
DR HSSP; P00451; 1CFG.
DR GO; GO:0005507; F: copper ion binding; IEA.
DR GO; GO:0007155; P: cell adhesion; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR000421; FA58_C.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00231; FA58C_2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS50022; FA58C_3; 2.
DR PROSITE; PS50079; MULTICOPPER OXIDASE1; 3.
DR PROSITE; PS50079; MULTICOPPER OXIDASE1; 3.
SQ SEQUENCE 2343 AA; 265613 MW; F612D744ADAADD99 CRC64;

Query Match
Best Local Similarity 83.9%; Score 731; DB 6; Length 2343;
Matches 133; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

QY 1 CDLNSCMPLGMSKAISSDAQITASSYFTNMATWSPSKARLHLQGRSNARWPQVNNPKE 60
DB 2180 CDLNSCMPLGMSKAISSDAQITASSYFTNMATWSPSKARLHLQGRSNARWPQVNNPKE 2239

QY 61 WLQVDFOKTMKVTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGNQD 120
DB 2240 WLQVDFOKTMKVTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFPQNGKVKVFGNQD 2299

QY 121 SFTPVNSLDPLLTRYLRIHPQSVWHQIALRMEVLGCEAQ 161
DB 2300 SFTPVNSLDPLLTRYLRIHPQSVWHQIALRMEVLGCEAQ 2340

RESULT 4
Q7TN96 PRELIMINARY; PRT; 2258 AA.
AC Q7TN96;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Factor VIII.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RA Watzka M., Geisen C., Seifried E., Oldenburg J.;
RT "Sequence of the rat factor VIII cDNA.";
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY362193; AAQ21580.1; -.
DR EMBL; AY362193; AAQ21580.1; -.
SQ SEQUENCE 2258 AA; 251299 MW; 13AF91C788059B1D CRC64;

```


Query Match 71.0%; Score 618; DB 11; Length 2258;
Best Local Similarity 68.7%; Pred. No. 4.5e-54;
Matches 112; Conservative 23; Mismatches 28; Indels 0; Gaps 0;

QY 1 CDLNSCMPGLMESKASDAQITASSYFTNNFATWSPSKARLHQGRSNARWPOVNNPKE 60
DB 2095 CDLNSCAPLGMESGWISDQVTFASSHLSTTFSAWPPALARHLRGANARWQVNDPTQ 2154

QY 61 WLQVDFQKTMKVTVGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFOGND 120
DB 2155 WLQVDFQKTMKVTVGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFOGND 2214

QY 121 SFTPVNSLDPLTRYLRIHQPSWVHQAIALRMEVLGCEAQLD 163
DB 2215 ASTPVNSLHPPTRYLRIHQPSWVHQAIALRMEVLGCEAQLD 2257

RESULT 5
ID Q804X3 PRELIMINARY; PRT; 1377 AA.
AC Q804X3;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Coagulation factor VIII precursor (Fragment).
GN F8.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Davidson C.J., Hirt R.P., Lal K., Snell P., Elgar G.,
RA Tuddenham E.G.D., McVey J.H.;
RT "Comparative sequence analysis and molecular evolution of blood
coagulation genes from Gallus gallus and Fugu rubripes.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF465272; AA033367.1; -
DR GO; GO:0005507; F:cell adhesion; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR000421; FA58 C.
DR Pfam; PF00394; Cu-oxidase; 1.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00022; FA58C_3; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
FT NON TER 1
SQ SEQUENCE 1377 AA; 152728 MW; 15ACFD7F7656565CD CRC64;

Query Match 65.2%; Score 568; DB 13; Length 1377;
Best Local Similarity 61.0%; Pred. No. 3.3e-49;
Matches 100; Conservative 30; Mismatches 34; Indels 0; Gaps 0;

QY 1 CDLNSCMPGLMESKASDAQITASSYFTNNFATWSPSKARLHQGRSNARWPOVNNPKE 60
DB 1214 CDLNSCMPGLMESKIGIPDQIRASYSYNNIFSSKNSQARLNLQGRTNARWPTNSPSE 1273

QY 61 WLQVDFQKTMKVTVGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFOGND 120
DB 1274 WLQVDFQKTMKVTVGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFOGND 1333

QY 121 SFTPVNSLDPLTRYLRIHQPSWVHQAIALRMEVLGCEAQLD 164
DB 1334 YTSVNSLSEPLFARYVRIHQPSWVHQAIALRMEVLGCEAQLD 1377

RESULT 6
ID Q804W6 PRELIMINARY; PRT; 1639 AA.
AC Q804W6;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Coagulation factor VIII precursor.
GN F8.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neoceratopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA Davidson C.J., Hirt R.P., Lal K., Snell P., Elgar G.,
RA Tuddenham E.G.D., McVey J.H.;
RT "Comparative sequence analysis and molecular evolution of blood
coagulation genes from Gallus gallus and Fugu rubripes.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF465279; AA033374.1; -
DR GO; GO:0005507; F:cell adhesion; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR000421; FA58 C.
DR InterPro; IPR008979; Gal Bind like.
DR Pfam; PF00394; Cu-oxidase; 2.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00231; FA58C; 2.
DR PROSITE; PS01286; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00022; FA58C_3; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
SQ SEQUENCE 1639 AA; 185203 MW; 881778A9F69006 CRC64;

Query Match 53.4%; Score 465; DB 13; Length 1639;
Best Local Similarity 50.9%; Pred. No. 1.6e-38;
Matches 83; Conservative 28; Mismatches 50; Indels 2; Gaps 1;

QY 1 CDLNSCMPGLMESKASDAQITASSYFTNNFATWSPSKARLHQGRSNARWPOVNNPKE 60
DB 1474 CDLNSCMPGLQDRRIPDESFASSYSWLSRWIPSLARLHQGRSNARWPKNNPHE 1533

QY 61 WLQVDFQKTMKVTVGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNG--KVKVFOGN 118
DB 1534 WLQVDFQKTMKVTVGTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNG--KVKVFOGN 1593

QY 119 QDSFTPVNSLDPLTRYLRIHQPSWVHQAIALRMEVLGCEAQLD 161
DB 1594 NDSDEALTFIDAPLFGYRIHQPSWVHQAIALRMEVLGCEAQLD 1636

RESULT 7
ID Q7SZNO PRELIMINARY; PRT; 1460 AA.
AC Q7SZNO;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Pseudarin C precursor.
OS Pseudonaja textilis (Eastern brown snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Acanthophiinae; Pseudonaja.
OX NCBI_TaxID=8673;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=22781770; PubMed=12730119;
RA Rao V.S., Swarup S., Kini R.M.;

"The nonenzymatic subunit of pseudotarin C, a prothrombin activator from eastern brown snake (*Pseudonaja textilis*) venom, shows structural similarity to mammalian coagulation factor V.";
 RL Blood 102:1347-1354(2003).
 DR EMBL; AY168281; AAC38805.1; -.
 KW SIGNAL.
 FT SIGNAL
 SQ SEQUENCE 1460 AA; 165931 MW; 6APB63E2D5D275A6 CRC64;
 1 10 Potential.
 Query Match 46.5%; Score 405; DB 13; Length 1460;
 Best Local Similarity 47.8%; Pred. No. 2.1e-32;
 Matches 77; Conservative 31; Mismatches 51; Indels 2; Gaps 1;
 QY 1 CDLNSCSNPLGMSKAISSAQITASSYFTNNFATWSPSKARHLQGRSNARWQVANNPK 60
 DB 1298 CEVGCSPVLMGSGATNASEITASSYKKTWSSWEPFLARLNLEGGTNWQVANNKQ 1357
 QY 61 WLQVDFQKTMKVGVTTQGVKSLTSMYKFLISSQDGHQWT-LFFQNGKVKVPOGN 118
 DB 1358 WLQIDLQHLTKITSIITQGGATSMYKVTFSHYTDDNSTWKPVLVDVRTSMKVFETGN 1417
 QY 119 QDSFTPVVNSLDPPLLRYLRIHPQSWHQAIALRMEVLGCE 159
 DB 1418 INSGVHKHFPKLPILSRFIRIPKTNQVIALRLELFGCE 1458
 RESULT 8
 ID Q804X4 PRELIMINARY; PRT; 745 AA.
 AC Q804X4;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Coagulation factor V precursor (Fragment).
 GN F5.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Davidson C.J., Hirt R.P., Lal K., Snell P., Elgar G.,
 RA Tuddenham E.G.D., McVey J.H.;
 RT "Comparative sequence analysis and molecular evolution of blood coagulation genes from Gallus gallus and Fugu rubripes.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF465271; AAC33366.1; -.
 DR GO; GO:0005507; F:copper ion binding; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR001117; Cu-oxidase.
 DR InterPro; IPR008972; Cupredoxin.
 DR InterPro; IPR008979; Gal Bind like.
 DR Pfam; PF00394; Cu-oxidase; 1.
 DR SMART; SM00231; FAS8C; 2.
 DR PROSITE; PS01285; FAS8C_1; 2.
 DR PROSITE; PS01286; FAS8C_2; 2.
 DR PROSITE; PS50022; FAS8C_3; 2.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.
 FT NON TER 1
 SQ SEQUENCE 745 AA; 85626 MW; A264587A348C29EE CRC64;
 Query Match 46.2%; Score 402; DB 13; Length 745;
 Best Local Similarity 44.7%; Pred. No. 1.8e-32;
 Matches 72; Conservative 40; Mismatches 47; Indels 2; Gaps 1;
 QY 1 CDLNSCSNPLGMSKAISSAQITASSYFTNNFATWSPSKARHLQGRSNARWQVANNPK 60
 DB 581 CEVDACSLPLGMENGEIKNTQITASSYFTWSPFLARINQKINAWRAESNNQO 640
 QY 61 WLQVDFQKTMKVGVTTQGVKSLTSMYKFLISSQDGHQWT-LFFQNGKVKVPOGN 118

DB 641 WLQIDLLIKKITAATQGVSTTENVKVIYLYNQSGSEWKSYTESSVAKVPSGN 700
 QY 119 QDSFTPVVNSLDPPLLRYLRIHPQSWHQAIALRMEVLGCE 159
 DB 701 EDSRGVHKHFPKLPILSRFIRIPKTNQVIALRLELFGCD 741
 RESULT 9
 ID O88783 PRELIMINARY; PRT; 2183 AA.
 AC O88783;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Murine coagulation factor V.
 GN F5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=9828202; PubMed=9616155;
 RA Yang T.L., Cui J., Rehumtulla A., Yang A., Moussalli M., Kaufman R.J.,
 RA Ginsburg D.;
 RT "The structure and function of murine factor V and its inactivation by blood protein C";
 RL Blood 91:4593-4599(1998).
 CC !- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 DR EMBL; U52925; AAC99553.1; -.
 DR PIR; T42764; T42764.
 DR HSSP; P12259; 1CZT.
 DR MGD; MGI:88382; F5.
 DR GO; GO:0005507; F:copper ion binding; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR001117; Cu-oxidase.
 DR InterPro; IPR008972; Cupredoxin.
 DR InterPro; IPR000421; FAS8_C.
 DR InterPro; IPR008979; Gal Bind like.
 DR Pfam; PF00394; Cu-oxidase; 3.
 DR Pfam; PF00754; F5 F8 type_C; 2.
 DR SMART; SM00231; FAS8C; 2.
 DR PROSITE; PS01285; FAS8C_1; 2.
 DR PROSITE; PS01286; FAS8C_2; 2.
 DR PROSITE; PS50022; FAS8C_3; 2.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
 SQ SEQUENCE 2183 AA; 247228 MW; BF0A8AA723F60317 CRC64;
 Query Match 44.8%; Score 390; DB 11; Length 2183;
 Best Local Similarity 45.5%; Pred. No. 1.2e-30;
 Matches 76; Conservative 31; Mismatches 54; Indels 6; Gaps 3;
 QY 1 CDLNSCSNPLGMSKAISSAQITASSYFTNNFATWSPSKARHLQGRSNARWQVANNPK 59
 DB 2020 CEVNGCSTPLGLEDGRIDQKQITASSFKKSWGWDYEPSLARLNAQGRVANAKANNK 2079
 QY 60 WLQVDFQKTMKVGVTTQGVKSLTSMYKFLISSQDGHQWT-LFFQNGKVKV--KVFQ 117
 DB 2080 QWLQVLLKKIKKTAIVTQGCSSLSEMYKSIQYSDQGVAKVPYRQKSMVDKIFEG 2139
 QY 118 NQDSFTPVVNSLDPPLLRYLRIHPQSWHQAIALRMEVLGCEAQDLY 164
 DB 2140 NSNTKGHKHFPKLPILSRFIRIPKTNQVIALRLELFGC---DIY 2183
 RESULT 10
 ID Q7TPK2 PRELIMINARY; PRT; 2102 AA.
 AC Q7TPK2;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```

DE Ac2-120.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu C.S., Li W.O., Li Y.C., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
RA Yang K.J., Wan H.M., Chang C.F., Zhao L.F., Ma H., Wang L., Wang S.F.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RT "Liver regeneration after PH.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY321333; AAP86265.1; -
SQ SEQUENCE 2102 AA; 237257 MW; 001BCB7DE31AC560 CRC64;

Query Match 44.2%; Score 385; DB 11; Length 2102;
Best Local Similarity 44.3%; Pred. No. 3.7e-30;
Matches 74; Conservative 34; Mismatches 53; Indels 6; Gaps 3;

QY 1 CDLNSCSMPGLMESKAISSDAQITASSYFTNMFAT-WSPSKARLHLQGRSNARPNQVNP 59
DB 1939 CEVNGCSTPLGLEDRIGNQKITASSFKKSWGSGYWEPSLARLNAQGRVNAWQAKANNK 1998

QY 60 EWLVDFQKTMKVTGVTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKV--KVFOG 117
DB 1999 QWLQIDLLKTKVTAIVTQGLCKLSSEMYVKSYSILYSDGVSWKPYRKSSMVDKIFEG 2058

QY 118 NQDSFTPVVNSLDPPLTRYLRHPQSWHQAIALRMEVLGCEAQDLY 164
DB 2059 NSNTKGHKMKFFNPPIISRIRIPIKTNQSIARLLELFGC---DIV 2102

RESULT 11
ID Q8N610 PRELIMINARY; PRT; 480 AA.
Q8N610:
AC Q8N610;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE EGF-like repeats and discoidin I-like domains 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RC Tissue=Brain;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC030828; AAH30828.1; -
DR GO; GO:0005509; P:calcium ion binding; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000421; FA58_C.
DR InterPro; IPR008579; Gal_Bind_like.
DR InterPro; IPR006210; IEGF.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00179; EGF_Ca; 3.
DR SMART; SM00231; FA58C; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_Ca; 1.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00022; FA58C_3; 2.
DR PROSITE; PS00022; EGF_1; 2.
KW EGF-like domain.
SQ SEQUENCE 480 AA; 53795 MW; 884E47ACFA507211 CRC64;

```

```

Query Match 42.1%; Score 366.5; DB 4; Length 480;
Best Local Similarity 46.4%; Pred. No. 4.5e-29;
Matches 78; Conservative 24; Mismatches 59; Indels 7; Gaps 3;

QY 1 CDLNSCSMPGLMESKAISSDAQITASSYFTNMFATWSPSKARLHLQGRSNARPNQVNP 56
DB 314 CELSGCSEPLGMSGHIQDYQITASSIFRTLNMDMF-TWEPKARLDKQKRVNWTSGHN 372

QY 57 NPKELQVDFQKTMKVTGVTQGVKSLTSMYKFLISSQDGHQWTLFFQNGKV 114
DB 373 DQSQWLQVLDLVPVKVITQGAQKDFGHQVFGSYKLAYSNDGEHRTVQDEKQKDKV 432

QY 115 FQGNQDSFTPVVNSLDPPLTRYLRHPQSWHQAIALRMEVLGCEAQD 162
DB 433 FQGNFDNDTHRNKVIDPPIYARHRIILFWISWYGRITLRSELLGCTEEB 480

RESULT 12
ID Q43854 PRELIMINARY; PRT; 480 AA.
Q43854:
AC Q43854; O43855;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Integrin-binding protein DEL1 precursor.
DE DEL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.
RC Tissue=Embryonic lung;
RX MEDLINE=98083109; PubMed=9420328;
RA Hidai C., Zupancic T.J., Penta K., Mikhail A., Kawana M.,
RA Quettermous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D.,
RA Quettermous R., Hogan B.L.M., Snodgrass R., Quettermous T.;
RT "Cloning and characterization of developmental endothelial locus-1: an
RT embryonic endothelial cell protein that binds the alphavbeta3 integrin
RT receptor.";
RL Genes Dev. 12:21-33(1998).
CC -!- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
CC INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS
CC FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN
CC REGULATION OF VASCULAR MORPHOGENESIS OF REMODELING IN EMBRYONIC
CC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q43854-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q43854-2; Sequence=VSP_050006; VSP_050007;
CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 F5/F8 TYPE C DOMAINS.
DR EMBL; U70312; AAC02648.1; -
DR EMBL; U70313; AAC02649.1; -
DR HSP; P12259; 1CZT.
DR Genew; HGNC:3173; EDIL3.
DR GO; GO:0005176; F:integrin binding; TAS.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000421; FA58_C.
DR InterPro; IPR008579; Gal_Bind_like.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00179; EGF_Ca; 1.
DR SMART; SM00231; FA58C; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; 2.

```


Db 2016 PWIQVELKEYKITGIVTQAKSGMGNMFVRSYILEYSEDGRWMKYTDDEQKLFQ 2075

Qy 118 NQSFPTVNSLQPELLTRYLRHPQSWVFQIALRMEVLGC 159

Db 2076 NTDNGQIKNYIIPFIRIPKQWKQKSVTRIELLGC 2117

RESULT 15

Q9WTS3 PRELIMINARY; PRT; 426 AA.

AC Q9WTS3, 1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Milk fat globule glycoprotein MFG-E8 S (Similar to milk fat globule-BGF factor 8 protein).

GN MFG-E8.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=Mammary gland;

RA MEDLINE=99120894; PubMed=9920772;

RA Oshima K., Aoki N., Negi M., Kishida M., Kitajima K., Matsuoka T.;

RT "Lactation-dependent expression of an mRNA splice variant with an exon

RT for a multiply O-glycosylated domain of mouse milk fat globule

RT glycoprotein MFG-E8.";

RL Biochem. Biophys. Res. Commun. 254:522-528(1999).

RN [2]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=NOD;

RA MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.

DR EMBL; AB025280; BAA76386.1; -

DR EMBL; BC003904; AAH03904.1; -

DR EMBL; BC003892; AAH03892.1; -

DR EMBL; AK089211; BAC40794.1; -

DR HSSP; P00740; 1EDM.

DR MGD; MGI:102768; Mfge8.

DR GO; GO:0007155; P:Cell adhesion; IEA.

DR InterPro; IPR001438; EGF-II.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR000421; FA58_C.

DR InterPro; IPR008979; Gal_bind_like.

DR InterPro; IPR006210; IEGF.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00754; F5_F8_type_C; 2.

DR PRINTS; PR00010; EGF_BLOOD.

DR SMART; SM00181; EGF; 2.

DR SMART; SM00231; FA58C; 2.

DR PROSITE; PS00022; EGF_1; 2.

DR PROSITE; PS01186; EGF_2; 2.

DR PROSITE; PS01285; FA58C_1; 2.

DR PROSITE; PS01286; FA58C_2; 2.

DR PROSITE; PS00022; FA58C_3; 2.

DR EGF-like domain.

SW EGF-like domain.

SQ SEQUENCE 426 AA; 47197 MW; B192F4A20629881A CRC64;

Query Match 41.0%; Score 357.5; DB 11; Length 426;

Best Local Similarity 43.5%; Pred. No. 3.2e-28;

Matches 70; Conservative 31; Mismatches 57; Indels 3; Gaps 2;

Qy 1 CDLNSCMPLGWSKASDAQITASSYPT--NMFA-TWSPSKARHLQGRNAPQVNN 57

Db 266 CELHGCSEPLGLKNNITPDSQMSASSYKTNLRAFGWYPHLGRLDNQKINAWTAQNS 325

Qy 58 PKEWLQVDFQTKMKTGVTGQVKSLLTSMVYKFLISSODGHQWTLFFONGKVKVFOG 117

Db 326 AKEWLQVDFQTKMKTGVTGQVKSLLTSMVYKFLISSODGHQWTLFFONGKVKVFOG 117

Qy 118 NQDSFTPVNSLQPELLTRYLRHPQSWVFQIALRMEVLGC 158

Db 386 NLDNNSHKKNIFEKPFMARYVRVLPVSWHNRITLRLELGC 426

Search completed: September 29, 2004, 16:39:45

Job time : 43 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: September 29, 2004, 16:35:58 ; Search time 15 seconds
(without alignments)
1051.694 Million cell updates/sec

Title: US-10-049-399A-1_COPY_2169_2332
Perfect score: 871
Sequence: 1 CDLNSCSMPGLGSKAISDA.....WVHQIALRMEVLGCEAQDLY 164

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	871	100.0	216	2 A44258	factor VIII-associ
2	871	100.0	2351	1 EZHU	coagulation factor
3	752	86.3	2319	2 A47004	coagulation factor
4	720	82.7	2133	2 T42763	coagulation factor
5	390	44.8	2183	2 T42764	coagulation factor
6	380	43.6	2224	1 KFHUS	coagulation factor
7	379	43.5	2211	1 KPBOS	coagulation factor
8	367.5	42.2	427	2 JC4915	agg protein precu
9	353.5	40.6	463	1 A36479	milk fat globule m
10	328.5	37.7	409	2 T11743	pp47 protein - pig
11	327.5	37.6	218	2 A47285	milk fat globule p
12	324.5	37.3	401	2 S65138	glycoprotein antig
13	324.5	37.3	427	2 S74211	PA5-6/7 protein pr
14	262	30.1	845	2 JC5256	adipocyte transcri
15	254.5	29.2	927	1 JQ0948	A5 antigen precurs
16	234.5	26.9	3133	2 S52093	hemocytin - silkw
17	221	25.4	719	2 S51739	transcription repr
18	207.5	23.8	737	2 T13149	hypothetical prote
19	206.5	23.7	737	2 T15615	hypothetical prote
20	193	22.2	913	2 A48280	receptor tyrosine
21	192	22.0	876	2 A49508	protein-tyrosine k
22	184	21.1	1283	2 T13799	neurexin IV - fru
23	183	21.0	910	2 A53137	tyrosine kinase re
24	176.5	20.3	855	2 S42621	protein-tyrosine k
25	171	19.6	1381	2 T31083	paranodin - rat
26	171	19.6	1385	2 T14158	neurexin IV - mous
27	156.5	18.0	819	2 I48859	tyro 10 receptor k
28	134.5	15.4	791	2 T16031	hypothetical prote
29	84.5	9.7	149	1 DLDOID	discooidin I chain

30 84.5 9.7 149 2 B03382
31 84.5 9.7 253 1 DLDOIC
32 82.5 9.5 253 1 DLDOIA
33 80.5 9.2 157 2 H69849
34 80.5 9.2 881 2 S67026
35 79 9.1 334 2 H75362
36 79 9.1 464 2 T01324
37 78.5 9.0 762 2 H71916
38 78 9.0 491 2 I40455
39 78 9.0 1134 2 T20332
40 77.5 8.9 1828 2 B59254
41 77.5 8.9 1853 1 A46761
42 77.5 8.9 1855 2 A59254
43 76 8.7 202 2 AC0168
44 76 8.7 551 2 C64894
45 75 8.6 589 2 T35784

ALIGNMENTS

RESULT 1
A44258
factor VIII-associated gene B hypothetical protein - human
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Jun-1999
C:Accession: A44258
R:Levinson, B.; Kenrick, S.; Gamel, P.; Fisher, K.; Gitschier, J.
Genomics 14, 585-589, 1992
A>Title: Evidence for a third transcript from the human factor VIII gene.
A:Reference number: A44258; MUID:93052386; PMID:1427887
A:Accession: A44258
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-216 <LEV>
A:Cross-references: GB:M30707; NID:g182316; PID:AAA58466.1; PID:g182317
C:Superfamily: coagulation factor VIII; discooidin I amino-terminal homology; ferroxiidas
F1-55/Domain: discooidin I amino-terminal homology #status atypical <DN1>
F157-210/Domain: discooidin I amino-terminal homology <DN2>

Query Match 100.0%; Score 871; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No.:1.3e-76;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CDLNSCSMPGLGSKAISDAQITASSYFTNNMFATWSPSKARLHLQGRSNAPRQVNNPKE 60
DB 53 CDLNSCSMPGLGSKAISDAQITASSYFTNNMFATWSPSKARLHLQGRSNAPRQVNNPKE 112
QY 61 WLQVDFQKTMKVTGVTGQVKSLLTSMYVKBFLISSQDGHQWTLFFQNGKVKVFGNQD 120
DB 113 WLQVDFQKTMKVTGVTGQVKSLLTSMYVKBFLISSQDGHQWTLFFQNGKVKVFGNQD 172
QY 121 SFTPVNSLDPPLLTRYLRIRHPQSWHQAIRMEVLGCEAQDLY 164
DB 173 SFTPVNSLDPPLLTRYLRIRHPQSWHQAIRMEVLGCEAQDLY 216

RESULT 2
EZHU
coagulation factor VIII precursor [validated] - human
N:Alternate names: antithrombolic factor A; coagulation factor VIIIc; procoagulant comp
C:Species: Homo sapiens (man)
C>Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Dec-2000
C:Accession: I54318; A00525; I58059; A23584; A26174; A42348; A43986; S66445; B4
R:Gitschier, J.; Wood, W.I.
Hum. Mol. Genet. 1, 199-200, 1992
A>Title: Sequence of the exon-containing regions of the human factor VIII gene.
A:Reference number: I54318; MUID:93265012; PMID:1303178
A:Accession: I54318
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1921, 'S', 1923-2351 <RES>
A:Cross-references: GB:M88648; NID:g182381; PID:AAA52420.1; PID:g182383

R;Wood, W.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschier, J.; Keyt, B.; Seeburg Nature 312, 330-337, 1984
 A>Title: Expression of active human factor VIII from recombinant DNA clones.
 A;Reference number: A00525; MUID:85061548; PMID:6438526
 A;Accession: A00525
 A;Molecule type: mRNA
 A;Residues: 1-2351 <MOO>
 A;Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179
 R;Toole, J.J.; Knopf, J.B.; Wozney, J.M.; Sultzman, L.A.; Pittman, D.D.; S, D.N.; Hewick, R.M.
 Nature 312, 342-347, 1984
 A>Title: Molecular cloning of a cDNA encoding human antihemophilic factor.
 A;Reference number: I58059; MUID:85061550; PMID:6438528
 A;Accession: I58059
 A;Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-74, 'V', 76-1259, 'E', 1261-2351 <RE2>
 A;Cross-references: GB:K01740; NID:G182802; PIDN:AAA52484.1; PID:G182803
 R;Truett, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dina, D.; Hartog, K.; Kuo, B.; Randolph, A.; Urdea, M.S.; Valenzuela, P.; Dahl, H.H.; Favalaro, J.; Hansen, J.; N DNA 4, 333-349, 1985
 A>Title: Characterization of the polypeptide composition of human factor VIII:C and the
 A;Reference number: A23584; MUID:86081164; PMID:3935400
 A;Accession: A23584
 A;Molecule type: protein
 A;Residues: 1-2351 <TRU>
 A;Cross-references: GB:M14113; NID:G182817; PIDN:AAA52485.1; PID:G182818
 R;Eaton, D.; Rodriguez, H.; Vehar, G.A.
 Biochemistry 25, 505-512, 1986
 A>Title: Proteolytic processing of human factor VIII. Correlation of specific cleavages
 ity.
 A;Reference number: A26174; MUID:86159740; PMID:3082357
 A;Accession: A26174
 A;Molecule type: protein
 A;Residues: 20-36;392-399, 'X', 401-402;1668-1678;1709-1722, 'D', 1723-1725;1741-1755 <EAT>
 R;Pittman, D.D.; Wang, J.H.; Kaufman, R.J.
 Biochemistry 31, 3315-3325, 1992
 A>Title: Identification and functional importance of tyrosine sulfate residues within re
 A;Reference number: A42348; MUID:92207952; PMID:1554716
 A;Accession: A42348
 A;Molecule type: protein
 A;Residues: 20-36;356-371;392-408;582-594;1668-1669, 'X', 1671;1672-1692;1693-1708;1709-17
 A;Experimental source: recombinant material from Chinese hamster ovary cells
 A;Note: sequence extracted from NCBI backbone and corrected to correspond with the publi
 R;Pay, P.J.; Smudzin, T.M.
 J. Biol. Chem. 264, 14005-14010, 1989
 A>Title: Intersubunit fluorescence energy transfer in human factor VIII.
 A;Reference number: A43986; MUID:89340500; PMID:2503509
 A;Accession: A43986
 A;Molecule type: protein
 A;Residues: 'X', 517-523;1853-1860, 'X', 1862-1864, 'X', 1866 <PAY>
 R;Keyte, A.; van Schijndel, H.B.; Niehrs, C.; Huttner, W.B.; Verbeet, M.P.; Mertens, K.; J. Biol. Chem. 266, 740-746, 1991
 A>Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for
 A;Reference number: A56109; MUID:91093266; PMID:1898735
 A;Contents: annotation; sulfation
 R;Gitschier, J.; Wood, W.I.; Goralka, T.M.; Wion, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar, Nature 312, 326-330, 1984
 A>Title: Characterization of the human factor VIII gene.
 A;Reference number: A56196; MUID:85061547; PMID:6438525
 A;Contents: annotation; introns
 R;McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.
 Protein Sci. 4, 740-746, 1995
 A>Title: Locations of disulfide bonds and free cysteines in the heavy and light chains o
 A;Reference number: A5216; MUID:95338127; PMID:7613471
 A;Contents: annotation; disulfide bonds
 R;Kjalke, M.; Heding, A.; Talbo, G.; Pettersen, E.; Thomsen, J.; Ezban, M.
 Eur. J. Biochem. 234, 773-779, 1995
 A>Title: Amino acid residues 721-729 are required for full factor VIII activity.
 A;Reference number: S63527; MUID:96163459; PMID:8575434
 A;Accession: S63527
 A;Molecule type: protein

A;Residues: 733-752;753-759 <XJA>
 R;Lind, P.; Larsson, K.; Spira, J.; Sydow-Baeckman, M.; Almstedt, A.; Gray, E.; Sandberg Eur. J. Biochem. 232, 19-27, 1995
 A>Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction
 A;Reference number: S66445; MUID:96048024; PMID:7556150
 A;Accession: S66445
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1669-1685 <LIN>
 C;Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure pr
 C;Genetics:
 A;Gene: GDB:F8C
 A;Cross-references: GDB:119124; OMIM:306700
 A;Map position: Xq28-Xq28
 A;Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63
 C;Function:
 A;Description: acts as a cofactor, with calcium and phospholipid, for the factor IXa pro
 A;Pathway: blood coagulation
 C;Superfamily: coagulation factor VIII; disocidin I amino-terminal homology; ferroxidase
 C;Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pia
 F;1-19/Domain: signal sequence #status predicted <SIG>
 F;20-2351/Product: coagulation factor VIII #status experimental <MAT>
 F;20-740/Product: coagulation factor VIIa heavy chain #status experimental <ACH>
 F;20-356/Domain: A1 <DA1>
 F;23-348/Domain: ferroxidase repeat homology <FO1>
 F;392-759/Domain: A2 <DA2>
 F;402-730/Domain: ferroxidase repeat homology <FO2>
 F;760-1667/Domain: B <DB0>
 F;1668-2351/Product: coagulation factor VIIa light chain #status experimental <ACL>
 F;1709-2038/Domain: A3 <DA3>
 F;2039-2191/Domain: C1 <DC1>
 F;2039-2188/Domain: disocidin I amino-terminal homology <DN1>
 F;2132-2351/Domain: C2 <DC2>
 F;2132-2345/Domain: disocidin I amino-terminal homology <DN2>
 F;602-258,601,776,803,847,919,962,1020,1024,1074,1085,1204,1274,1278,1301,1319,1403,1
 F;172-198,267-348,547-573,649-730,1851-1877,1918-1922,2040-2188/Disulfide bonds: #status
 F;355-356/Cleavage site: Arg-Met (coagulation factor Xa, protein C) #status predicted
 F;365,737,738,742,1683,1699/Binding site: sulfate (Tyr) (covalent) #status experimental
 F;391-392/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
 F;414,426/Binding site: sulfate (Tyr) (covalent) #status predicted
 F;759-760/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
 F;1667-1668/Cleavage site: Arg-Glu (unidentified proteinase) #status experimental
 F;1708-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
 F;1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) #status experimental
 F;2193-2345/Disulfide bonds: #status predicted
 Query Match 100.0%; Score 871; DB 1; Length 2351;
 Best Local Similarity 100.0%; Pred. No. 2.6e-75;
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CDLNSCMP;GMESKALSDAQITASSYFTNMPATWSPSKARLHLOGRSNARPOVNNPKE 60
 Db 2188 CDLNSCMP;GMESKALSDAQITASSYFTNMPATWSPSKARLHLOGRSNARPOVNNPKE 2247
 Qy 61 WLQVDFOKTKVTKVTQTQGVKSLTTSYKFEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
 Db 2248 WLQVDFOKTKVTKVTQTQGVKSLTTSYKFEFLISSQDGHQWTLFFQNGKVKVFGNQD 2307
 Qy 121 SFTFVNVSLDPPLLTRYLRTHPSQSWHQIALRMEVLGCEAQDLY 164
 Db 2308 SFTFVNVSLDPPLLTRYLRTHPSQSWHQIALRMEVLGCEAQDLY 2351
 RESULT 3
 A47004
 coagulation factor VIII precursor - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Jun-1999
 C;Accession: A47004
 R;Eider, B.; Lakich, D.; Gitschier, J.
 Genomics 16, 374-379, 1993
 A>Title: Sequence of the murine factor VIII cDNA.

A;Reference number: A47004; MUID:93300511; PMID:8314577
A;Accession: A47004
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2319 <ELD>
A;Cross-references: GB:L05573; NID:g192456; PIDN:AAA37385.1; PID:g192457
C;Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;19-349/Domain: ferroxidase repeat homology <FOI>
F;402-730/Domain: ferroxidase repeat homology <FO2>
F;1686-2006/Domain: ferroxidase repeat homology <FO3>
F;2007-2156/Domain: discoidin I amino-terminal homology <DN1>
F;2160-2313/Domain: discoidin I amino-terminal homology <DN2>

Query Match 86.3%; Score 752; DB 2; Length 2319;
Best Local Similarity 84.1%; Pred. No. 7.7e-64;
Matches 138; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

QY 1 CDLNSCSMPGLMGESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARWPQVNNPK 60
Db 2156 CDLNSCSIPGLMGESKVISDTQITASSYFTNMFATWSPSKARLHLQGRTNARWPQVNDPK 2215

QY 61 WLQVDFQKTMKVTGVTQGVKSLTSMYKFEFLISSQDGHQWTLFPQNGKVKVFGNQD 120
Db 2216 WLQVDFQKTMKVTGVTQGVKSLTSMYKFEFLISSQDGHQWTLFPQNGKVKVFGNQD 2275

QY 121 SFTPVVNSLDPPLTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 164
Db 2276 SSTPMVNSLDPPLTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 2319

RESULT 4
T42763
coagulation factor VIII precursor - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
C;Accession: T42763
R;Lollar, P.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z22269
A;Accession: T42763
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-2133 <LOL>
A;Cross-references: EMBL:U49517; NID:g1511633; PID:g1511634; PIDN:AAB06705.1
C;Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
C;Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-2133/Product: coagulation factor VIII #status predicted <MAT>
F;23-349/Domain: ferroxidase repeat homology <FOX1>
F;402-730/Domain: ferroxidase repeat homology <FOX2>
F;1498-1820/Domain: ferroxidase repeat homology <FOX3>

Query Match 92.7%; Score 720; DB 2; Length 2133;
Best Local Similarity 79.9%; Pred. No. 8.5e-61;
Matches 131; Conservative 20; Mismatches 13; Indels 0; Gaps 0;

QY 1 CDLNSCSMPGLMGESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARWPQVNNPK 60
Db 1970 CDLNSCSMPGLMGESKAISDAQITASSYFTNMFATWSPSKARLHLQGRTNARWPQVNNPK 2029

QY 61 WLQVDFQKTMKVTGVTQGVKSLTSMYKFEFLISSQDGHQWTLFPQNGKVKVFGNQD 120
Db 2030 WLQVDFQKTMKVTGVTQGVKSLTSMYKFEFLISSQDGHQWTLFPQNGKVKVFGNQD 2089

QY 121 SFTPVVNSLDPPLTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 164
Db 2090 SSTPMVNSLDPPLTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 2133

RESULT 5
T42764
coagulation factor V - mouse

C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
C;Accession: T42764
R;Yang, T.L.; Cui, J.; Rehmtulla, A.; Yang, A.; Mousallali, M.; Kaufman, R.J.; Ginsburg
Blood 91, 4593-4599, 1998
A;Title: The structure and function of murine factor V and its inactivation by protein
A;Reference number: Z22270; MUID:9828202; PMID:9616155
A;Accession: T42764
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-2183 <YAN>
A;Cross-references: EMBL:U52925; NID:g3219690; PID:g3219691; PIDN:AAC99553.1
C;Function:
A;Pathway: blood coagulation
C;Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase
C;Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasma;
F;350-682/Domain: ferroxidase repeat homology <FOX1>
F;1541-1864/Domain: ferroxidase repeat homology <FOX2>

Query Match 44.8%; Score 390; DB 2; Length 2183;
Best Local Similarity 45.5%; Pred. No. 6.1e-29;
Matches 76; Conservative 31; Mismatches 54; Indels 6; Gaps 3;

QY 1 CDLNSCSMPGLMGESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARWPQVNNPK 59
Db 2020 CEVAGCSTPLGLEDRIQDKQITASSSPKSKWGWYSEPLARLNAQGRVNAWQAKANNK 2079

QY 60 EWLQVDFQKTMKVTGVTQGVKSLTSMYKFEFLISSQDGHQWTLFPQNGKVKVFGNQD 117
Db 2080 QWLQVDFQKTMKVTGVTQGVKSLTSMYKFEFLISSQDGHQWTLFPQNGKVKVFGNQD 2139

QY 118 NODSFTPVVNSLDPPLTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 164
Db 2140 NSNTKGHMKNFNPPIISFRIRIIPKTNQSIARLELFGC---DIY 2183

RESULT 6
KFH5
coagulation factor V precursor [validated] - human
N;Alternate names: coagulation labile factor; proaccelerin
C;Species: Homo sapiens (man)
C;Date: 19-May-1989 #sequence_revision 02-Jun-1995 #text_change 08-Dec-2000
C;Accession: A56172; A42344; A28028; A27498; A25897
R;Cripe, L.D.; Moore, K.D.; Kane, W.H.
Biochemistry 31, 3777-3785, 1992
A;Title: Structure of the gene for human coagulation factor V.
A;Reference number: A42344; MUID:92232668; PMID:1567832
A;Accession: A56172
A;Molecule type: DNA
A;Residues: 1-2224 <CR1>
A;Cross-references: GB:J05368
A;Accession: A42344
A;Molecule type: DNA
A;Residues: 48-58; 79-89; 120-130; 191-201; 239-249; 313-323; 368-378; 428-437; 461-471; 533-542
2070; 2111-2120; 2172-2181 <CR2>
R;Jenny, R.J.; Pittman, D.D.; Toole, J.J.; Kriz, R.W.; Aldape, R.A.; Hewick, R.M.; Kauf
proc. Natl. Acad. Sci. U.S.A. 84, 4846-4850, 1987
A;Title: Complete cDNA and derived amino acid sequence of human factor V.
A;Reference number: A28028; MUID:87260886; PMID:3110773
A;Accession: A28028
A;Molecule type: mRNA
A;Residues: 1-857; 'R', 856-864; 'E', 926-924; 'I', 1765-2212; 'T', 2214-2224 <JEN
A;Cross-references: GB:M16967
A;Note: parts of this sequence, including the amino end of the mature protein, were det
R;Kane, W.H.; Ichinose, A.; Hagen, F.S.; Davie, E.W.
Biochemistry 26, 6508-6514, 1987
A;Title: Cloning of cDNAs coding for the heavy chain region and connecting region of hu
A;Reference number: A27498; MUID:88107560; PMID:2827731
A;Accession: A27498
A;Molecule type: mRNA
A;Residues: 1-1284; 'I', 1286-1600 <KAN>
A;Cross-references: GB:M17785
A;Note: parts of this sequence were determined by protein sequencing

QY 1 CDLNSCMPGLGMSKALSDAQITASSYFT--NWFA-TWSPSKARLHLQGRSNARPOVNN 57
 Db 303 CEHGCELEPLGLKNWTPISQMSASSYKTNWLRAGFNYPHLGRLDNQGKINAWTAQNS 362
 QY 58 PKEWLQVDFOKTKMKTGVTTQGVKSLTTSNYKVEFLISSQDGHQWTLFPQNGKVKVFOG 117
 Db 363 AKEWLQVDLQTORQVGTIITQGARDFGHIQYVESYKVAHSDGQVQWTVVEQSSKVFQ 422
 QY 118 NQDSFTPVNLSLDPPLTRYLRTHPOSVHQIALRMEVLGC 158
 Db 423 NLDNNSHKNIPEKPMARYVRVLPVSWENRITLRLELGC 463
 RESULT 10
 T11743
 PP47 protein - pig (fragment)
 C/Species: Sus scrofa domestica (domestic pig)
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 08-Sep-2002
 C/Accession: T11743
 R/Eusslin, M.; Vogel, T.; Calvete, J.J.; Thole, H.H.; Schmidtke, J.; Matsuda, T.; Toepf
 Biol. Reprod. 58, 1057-1064, 1998
 A/Title: Molecular cloning and characterization of P47, a novel boar sperm-associated z
 A/Reference number: Z17325; MUID:98206817; PMID:9546740
 A/Accession: T11743
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: mRNA
 A/Residues: 1-409 <ENS>
 A/Cross-references: EMBL:Y11693; NID:G2652927; PIDN:CAA72379.1; PID:G2652928
 A/Experimental source: testis
 C/Function:
 A/Description: may be involved in membrane remodeling and/or function as a zona pelluci
 C/Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homol
 F:6-40/Domain: EGF homology <EGF>
 Query Match 37.78; Score 328.5; DB 2; Length 409;
 Best Local Similarity 39.98; Pred. No. 6.4e-24;
 Matches 65; Conservative 31; Mismatches 62; Indels 5; Gaps 2;
 QY 1 CDLNSCMPGLGMSKALSDAQITASSYFTN--MFATWSPSKARLHLQGRSNARPOVNN 57
 Db 247 CELSGCAEPLGLKNDTIPNKQITASSPYRTWGLSAFWSYPFYARLDNQGKFNAWTAQNS 306
 QY 58 PKEWLQVDFOKTKMKTGVTTQGVKSLTTSNYKVEFLISSQDGHQWTLFPQNGKV--KVF 115
 Db 307 ASEWLQVDLQSGRRVVTGIITCGARDFGHIQYVAIKVAYSDGVSFTEYRDDQALEGKIF 366
 QY 116 QGNQDSFTPVNLSLDPPLTRYLRTHPOSVHQIALRMEVLGC 158
 Db 367 PGNLDNNSHKKNPFETPLTRFVRILPVAWNRIITLRVELLGC 409
 RESULT 11
 A47285
 milk fat globule protein - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 28-May-1999
 C/Accession: A47285
 R/Latocca, D.; Peterson, J.A.; Urrea, R.; Kuniyoshi, J.; Bistrain, A.M.; Ceriani, R.L.
 Cancer Res. 51, 4594-4599, 1991
 A/Title: A M-r 46,000 human milk fat globule protein that is highly expressed in human
 A/Reference number: A47285; MUID:91371351; PMID:1909932
 A/Accession: A47285
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-218 <LAR>
 A/Cross-references: GB:856151; NID:G235396; PIDN:RAB1977.1; PID:G235397
 C/Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homol
 F:1-56/Domain: discoidin I amino-terminal homology (fragment) <DN1>
 F:160-218/Domain: discoidin I amino-terminal homology <DN2>
 Query Match 37.68; Score 327.5; DB 2; Length 218;
 Best Local Similarity 40.28; Pred. No. 3.7e-24;
 Matches 66; Conservative 35; Mismatches 56; Indels 7; Gaps 3;

Job time : 16 secs

RESULT 15

Query Match	29.2%	Score 254.5;	DB 1;	Length 927;
Best Local Similarity	34.6%;	Pred. No. 2.5e-16;		
Matches	55;	Conservative	26;	Mismatches 71;
				Indels 7;
				Gaps 3;

Search completed: September 29, 2004, 16:40:24

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 16:38:04 ; Search time 33 Seconds
(without alignments)
256.565 Million cell updates/sec

Title: US-10-049-399a-1_COPY_2169_2332
Perfect score: 871
Sequence: 1 CDLNSCSPLGMSKAISDA.....WHQIALRMEVLGCEAQLDLY 164

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/2/iaa/5A COMB.pcp.*
- 2: /cgn2_6/prodata/2/iaa/5B COMB.pcp.*
- 3: /cgn2_6/prodata/2/iaa/6A COMB.pcp.*
- 4: /cgn2_6/prodata/2/iaa/6B COMB.pcp.*
- 5: /cgn2_6/prodata/2/iaa/PCITUS COMB.pcp.*
- 6: /cgn2_6/prodata/2/iaa/backfile1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	871	100.0	1438	4	US-09-209-916-1 Sequence 1, Appli
2	871	100.0	1471	1	US-08-683-839B-3 Sequence 3, Appli
3	871	100.0	1661	2	US-08-882-083-2 Sequence 2, Appli
4	871	100.0	1661	2	US-08-558-107-2 Sequence 2, Appli
5	871	100.0	1661	3	US-09-243-539-2 Sequence 2, Appli
6	871	100.0	2332	1	US-07-864-004B-4 Sequence 4, Appli
7	871	100.0	2332	1	US-08-251-937A-4 Sequence 4, Appli
8	871	100.0	2332	1	US-08-212-133A-2 Sequence 2, Appli
9	871	100.0	2332	1	US-08-276-594A-2 Sequence 2, Appli
10	871	100.0	2332	1	US-08-474-503-2 Sequence 2, Appli
11	871	100.0	2332	2	US-08-670-707A-2 Sequence 2, Appli
12	871	100.0	2332	3	US-09-037-601-2 Sequence 3, Appli
13	871	100.0	2332	3	US-09-324-867-3 Sequence 3, Appli
14	871	100.0	2332	4	US-09-315-179-2 Sequence 2, Appli
15	871	100.0	2332	4	US-09-523-656-2 Sequence 2, Appli
16	871	100.0	2332	5	PCT-US93-03275-4 Sequence 4, Appli
17	871	100.0	2332	5	PCT-US94-13200-2 Sequence 2, Appli
18	871	100.0	2351	1	US-08-121-202-2 Sequence 2, Appli
19	871	100.0	2351	1	US-08-366-851A-2 Sequence 4, Appli
20	871	100.0	2351	4	US-10-133-907-4 Sequence 4, Appli
21	871	100.0	2351	6	5171844-2 Patent No. 5171844
22	867	99.5	218	1	US-07-607-538C-5 Sequence 5, Appli
23	867	99.5	218	2	US-08-182-402B-5 Sequence 5, Appli
24	867	99.5	218	4	US-09-364-185-5 Sequence 5, Appli
25	862	99.0	2351	6	5422260-1 Patent No. 5422260
26	846	97.1	160	2	US-08-162-402B-15 Sequence 15, Appli
27	846	97.1	160	4	US-09-331-793-9 Sequence 9, Appli

28	752	86.3	2304	3	US-09-324-867-4 Sequence 4, Appli
29	752	86.3	2319	1	US-08-212-133A-8 Sequence 8, Appli
30	752	86.3	2319	1	US-08-474-503-6 Sequence 6, Appli
31	752	86.3	2319	2	US-08-670-707A-6 Sequence 6, Appli
32	752	86.3	2319	3	US-09-037-601-6 Sequence 6, Appli
33	752	86.3	2319	4	US-09-315-179-6 Sequence 28, Appli
34	752	86.3	2319	4	US-09-523-656-28 Sequence 6, Appli
35	752	86.3	2319	5	PCT-US94-13200-6 Sequence 2, Appli
36	736	84.5	2343	3	US-09-324-867-2 Sequence 39, Appli
37	720	82.7	1443	2	US-08-670-707A-39 Sequence 39, Appli
38	720	82.7	1443	3	US-09-037-601-39 Sequence 39, Appli
39	720	82.7	1443	4	US-09-315-179-39 Sequence 38, Appli
40	720	82.7	1467	4	US-09-523-656-38 Sequence 38, Appli
41	720	82.7	2115	3	US-09-324-867-5 Sequence 5, Appli
42	720	82.7	2133	2	US-08-670-707A-37 Sequence 37, Appli
43	720	82.7	2133	3	US-09-037-601-37 Sequence 37, Appli
44	720	82.7	2133	4	US-09-315-179-37 Sequence 37, Appli
45	720	82.7	2133	4	US-09-523-656-30 Sequence 30, Appli

ALIGNMENTS

RESULT 1
US-09-209-916-1
; Sequence 1, Application US/09209916
; Patent No. 6358703
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255
; CURRENT APPLICATION NUMBER: US/09/209,916
; CURRENT FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-09-209-916-1

Query Match	100.0%	Score	871	DB	4	Length	1438
Best Local Similarity	100.0%	Pred. No.	4e-94				
Mismatches	164	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	CDLNSCSPLGMSKAISDAQITASSYFTNMFATWSPSKARHLQGRNMRPQVNNPE	60				
Db	1275	CDLNSCSPLGMSKAISDAQITASSYFTNMFATWSPSKARHLQGRNMRPQVNNPE	1334				
QY	61	WLQVDFQTKMTGVTGTVKSLTSMYVKEFLISSSDQGHQWTLFFQNGKVKVFGNQD	120				
Db	1335	WLQVDFQTKMTGVTGTVKSLTSMYVKEFLISSSDQGHQWTLFFQNGKVKVFGNQD	1394				
QY	121	SFTPVVNSLDPPLLRYLRIHQSWHQAIALRMEVLGCEAQLDLY	164				
Db	1395	SFTPVVNSLDPPLLRYLRIHQSWHQAIALRMEVLGCEAQLDLY	1438				

RESULT 2
US-08-683-839B-3
; Sequence 3, Application US/08683839B
; Patent No. 5744326
; GENERAL INFORMATION:
; APPLICANT: Iil, Charles . R. et al.
; TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional
; TITLE OF INVENTION: Regulatory Sequences To Increase Expression of
; TITLE OF INVENTION: Introns Genes Containing Near-Consensus Splice Sites

; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street, suite 510
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109-1875
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/683,839B
 ; FILING DATE: 11-MARCH-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; NAME: Remillard, Jane B.
 ; REGISTRATION NUMBER: 38,872
 ; REFERENCE/DOCKET NUMBER: TTI-138
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617)227-7400
 ; TELEFAX: (617)227-5941
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1471 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-683-839B-3

Query Match 100.0%; Score 871; DB 1; Length 1471;
 Best Local Similarity 100.0%; Pred. No. 4.1e-94;
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CDLNCSMPGLGMSKAIISDAQITASSYFTNMFTWSPSKARLHLQGRSNARPPQVNNPKE 60
 DB 1308 CDLNCSMPGLGMSKAIISDAQITASSYFTNMFTWSPSKARLHLQGRSNARPPQVNNPKE 1367
 QY 61 WLQVDFQKTKVGTGVTGGVKSLLTSYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
 DB 1368 WLQVDFQKTKVGTGVTGGVKSLLTSYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 1427
 QY 121 SFTPVNSLDPPLLTTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 164
 DB 1428 SFTPVNSLDPPLLTTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 1471

RESULT 3
 US-08-682-083-2
 ; Sequence 2, Application US/0882083
 ; Patent No. 5869292
 ; GENERAL INFORMATION:
 ; APPLICANT: VOORBERG, Johannes J.
 ; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/882,083
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/558,107
 ; FILING DATE: 13-NOV-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ISACSON, John P.
 ; REGISTRATION NUMBER: 33,715
 ; REFERENCE/DOCKET NUMBER: 30472/212
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)672-5300
 ; TELEFAX: (202)672-5399
 ; TELEX: 904136
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1661 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-882-083-2

Query Match 100.0%; Score 871; DB 2; Length 1661;
 Best Local Similarity 100.0%; Pred. No. 4.9e-94;
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CDLNCSMPGLGMSKAIISDAQITASSYFTNMFTWSPSKARLHLQGRSNARPPQVNNPKE 60
 DB 1498 CDLNCSMPGLGMSKAIISDAQITASSYFTNMFTWSPSKARLHLQGRSNARPPQVNNPKE 1557
 QY 61 WLQVDFQKTKVGTGVTGGVKSLLTSYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
 DB 1558 WLQVDFQKTKVGTGVTGGVKSLLTSYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 1617
 QY 121 SFTPVNSLDPPLLTTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 164
 DB 1618 SFTPVNSLDPPLLTTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 1661

RESULT 4
 US-08-558-107-2
 ; Sequence 2, Application US/08558107
 ; Patent No. 5910481
 ; GENERAL INFORMATION:
 ; APPLICANT: VOORBERG, Johannes J.
 ; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/558,107
 ; FILING DATE: 13-NOV-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ISACSON, John P.
 ; REGISTRATION NUMBER: 33,715
 ; REFERENCE/DOCKET NUMBER: 30472/212
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)672-5300
 ; TELEFAX: (202)672-5399
 ; TELEX: 904136
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:

LENGTH: 1661 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-558-107-2

Query Match 100.0%; Score 871; DB 2; Length 1661;
Best Local Similarity 100.0%; Pred. No. 4.9e-94;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCSMPLGWSKASDAQITASSYFTNMFAWSPSKARHLQGRSNARWPQVNNPKE 60
DB 1498 CDLNSCSMPLGWSKASDAQITASSYFTNMFAWSPSKARHLQGRSNARWPQVNNPKE 1557
QY 61 WLQVDFQKTMKVTGVTGQVKSLLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGQND 120
DB 1558 WLQVDFQKTMKVTGVTGQVKSLLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGQND 1617
QY 121 SFTPVVNSLDPPLRLRYLRHQPQSWHQAIALRMEVLGCEAODLY 164
DB 1618 SFTPVVNSLDPPLRLRYLRHQPQSWHQAIALRMEVLGCEAODLY 1661

RESULT 5

US-09-243-539-2
Sequence 2, Application US/09243539
Patent No. 6130203
GENERAL INFORMATION:
APPLICANT: VOORBERG, Johannes J.
TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
NUMBER OF SEQUENCES: 17
CURRENT APPLICATION DATA:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/243,539
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/558,107
FILING DATE: 13-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: ISACSON, John F.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 30472/212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-243-539-2

Query Match 100.0%; Score 871; DB 3; Length 1661;
Best Local Similarity 100.0%; Pred. No. 4.9e-94;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCSMPLGWSKASDAQITASSYFTNMFAWSPSKARHLQGRSNARWPQVNNPKE 60
DB 1498 CDLNSCSMPLGWSKASDAQITASSYFTNMFAWSPSKARHLQGRSNARWPQVNNPKE 1557

QY 61 WLQVDFQKTMKVTGVTGQVKSLLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGQND 120
DB 1558 WLQVDFQKTMKVTGVTGQVKSLLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGQND 1617
QY 121 SFTPVVNSLDPPLRLRYLRHQPQSWHQAIALRMEVLGCEAODLY 164
DB 1618 SFTPVVNSLDPPLRLRYLRHQPQSWHQAIALRMEVLGCEAODLY 1661

RESULT 6

US-07-864-004B-4
Sequence 4, Application US/07864004B
Patent No. 5364771
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
NUMBER OF SEQUENCES: 6
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/864,004B
FILING DATE: 07 APRIL 1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EMU106
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6508
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Liver CDNA sequence
US-07-864-004B-4

Query Match 100.0%; Score 871; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 8.2e-94;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCSMPLGWSKASDAQITASSYFTNMFAWSPSKARHLQGRSNARWPQVNNPKE 60
DB 2169 CDLNSCSMPLGWSKASDAQITASSYFTNMFAWSPSKARHLQGRSNARWPQVNNPKE 2228

QY 61 WLQVDFQKTMKVTGVTGQVKSLLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGQND 120
DB 2229 WLQVDFQKTMKVTGVTGQVKSLLTSMYKFEFLISSQDGHQWTLFFQNGKVKVFGQND 2288

QY 121 SFTPVVNSLDPPLRLRYLRHQPQSWHQAIALRMEVLGCEAODLY 164
DB 2289 SFTPVVNSLDPPLRLRYLRHQPQSWHQAIALRMEVLGCEAODLY 2332

```

RESULT 7
US-08-251-937A-4
; Sequence 4, Application US/08251937A
; Patent No. 5583209
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marshall S.
; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/251,937A
; FILING DATE: 31-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6367
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver CDNA sequence
US-08-251-937A-4

Query Match 100.0%; Score 871; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 8.2e-94;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARPQVNNPK 60
DB 2169 CDLNSCMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARPQVNNPK 2228
QY 61 WLQVDFQKTKMKTGVTTQGVKSLTSMYKVELISSQDGHQWTLFFQNGKVKVFGNQD 120
DB 2229 WLQVDFQKTKMKTGVTTQGVKSLTSMYKVELISSQDGHQWTLFFQNGKVKVFGNQD 2288
QY 121 SFTPVVNSLDPLLTLYLRIHPQSWWHQIALRMEVLGCEAQDLY 164
DB 2289 SFTPVVNSLDPLLTLYLRIHPQSWWHQIALRMEVLGCEAQDLY 2332

RESULT 8
US-08-212-133A-2
; Sequence 2, Application US/08212133A
; Patent No. 5663060
; GENERAL INFORMATION:

```

```

; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marshall S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,133A
; FILING DATE: March 11, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU/76677
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver CDNA sequence
US-08-212-133A-2

Query Match 100.0%; Score 871; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 8.2e-94;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARPQVNNPK 60
DB 2169 CDLNSCMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNARPQVNNPK 2228
QY 61 WLQVDFQKTKMKTGVTTQGVKSLTSMYKVELISSQDGHQWTLFFQNGKVKVFGNQD 120
DB 2229 WLQVDFQKTKMKTGVTTQGVKSLTSMYKVELISSQDGHQWTLFFQNGKVKVFGNQD 2288
QY 121 SFTPVVNSLDPLLTLYLRIHPQSWWHQIALRMEVLGCEAQDLY 164
DB 2289 SFTPVVNSLDPLLTLYLRIHPQSWWHQIALRMEVLGCEAQDLY 2332

RESULT 9
US-08-276-594A-2
; Sequence 2, Application US/08276594A
; Patent No. 5693499
; GENERAL INFORMATION:
; APPLICANT: YONEMURA, Hiroshi
; APPLICANT: TAJIMA, Yoshitaka
; APPLICANT: SUGAWARA, Kenishin
; APPLICANT: MASUDA, Kenichi
; TITLE OF INVENTION: PROCESS FOR PREPARING HUMAN COAGULATION
; TISSUE TYPE: FACTOR VIII PROTEIN COMPLEX

```

```

;
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,594A
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/950,191
; FILING DATE: 24-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 243262/1991
; FILING DATE: 24-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 74129/195/AOPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-276-594A-2

Query Match 100.0%; Score 871; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 8.2e-94;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCMPGLGMSKATSDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKE 60
Db 2169 CDLNSCMPGLGMSKATSDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKE 2228

QY 61 WLQVDFQKTMKVTGVTGQVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 2229 WLQVDFQKTMKVTGVTGQVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 2288

QY 121 SFTPVNSLDPPLTRYLRIRHPQSWHQAIALRMEVLGCEAODLY 164
Db 2289 SFTPVNSLDPPLTRYLRIRHPQSWHQAIALRMEVLGCEAODLY 2332

RESULT 10
US-08-474-503-2
; Sequence 2, Application US/08474503
; Patent No. 5744446
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
;
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,594A
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/950,191
; FILING DATE: 24-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 243262/1991
; FILING DATE: 24-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 74129/195/AOPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-276-594A-2

Query Match 100.0%; Score 871; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 8.2e-94;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCMPGLGMSKATSDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKE 60
Db 2169 CDLNSCMPGLGMSKATSDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKE 2228

QY 61 WLQVDFQKTMKVTGVTGQVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 2229 WLQVDFQKTMKVTGVTGQVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 2288

QY 121 SFTPVNSLDPPLTRYLRIRHPQSWHQAIALRMEVLGCEAODLY 164
Db 2289 SFTPVNSLDPPLTRYLRIRHPQSWHQAIALRMEVLGCEAODLY 2332

RESULT 11
US-08-670-707A-2
; Sequence 2, Application US/08670707A
; Patent No. 5859204
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,707A
; FILING DATE: 26-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US94/13200
; FILING DATE: 15-NOV-1994
;
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,503
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106CIP(3)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6500
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cDNA sequence
; US-08-474-503-2

Query Match 100.0%; Score 871; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 8.2e-94;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCMPGLGMSKATSDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKE 60
Db 2169 CDLNSCMPGLGMSKATSDAQITASSYFTNMFATWSPSKARLHLQGRSNARPOVNNPKE 2228

QY 61 WLQVDFQKTMKVTGVTGQVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 2229 WLQVDFQKTMKVTGVTGQVKSLLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 2288

QY 121 SFTPVNSLDPPLTRYLRIRHPQSWHQAIALRMEVLGCEAODLY 164
Db 2289 SFTPVNSLDPPLTRYLRIRHPQSWHQAIALRMEVLGCEAODLY 2332
```

```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,133
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
; US-08-670-707A-2

Query Match 100.0%; Score 871; DB 2; Length 2332;
Best Local Similarity 100.0%; Pred. No. 8.2e-94;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCMPGMSKALSDAQITASSYFTNMFWATWSPSKARHLQGRSNARPPQVNNPKE 60
Db 2169 CDLNSCMPGMSKALSDAQITASSYFTNMFWATWSPSKARHLQGRSNARPPQVNNPKE 2228

QY 61 WLQVDFQKTKVGTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 2229 WLQVDFQKTKVGTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQD 2288

QY 121 SFTPVNSLDPPLLTRYLRHPSQSWHQAIALRMEVLGCEAQDLY 164
Db 2289 SFTPVNSLDPPLLTRYLRHPSQSWHQAIALRMEVLGCEAQDLY 2332

RESULT 12
US-09-037-601-2
; Sequence 2, Application US/09037601
; Patent No. 6180371
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/037,601
; FILING DATE: 26-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US94/13200
; FILING DATE: 15-NOV-1994
; PRIOR APPLICATION DATA:
```

```
;
; APPLICATION NUMBER: US 08/212,133
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 75-95F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
; US-09-037-601-2

Query Match 100.0%; Score 871; DB 3; Length 2332;
Best Local Similarity 100.0%; Pred. No. 8.2e-94;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CDLNSCMPGMSKALSDAQITASSYFTNMFWATWSPSKARHLQGRSNARPPQVNNPKE 60
Db 2169 CDLNSCMPGMSKALSDAQITASSYFTNMFWATWSPSKARHLQGRSNARPPQVNNPKE 2228

QY 61 WLQVDFQKTKVGTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 2229 WLQVDFQKTKVGTGTTQGVKSLTSMYKVEFLISSQDGHQWTLFFQNGKVKVFGNQD 2288

QY 121 SFTPVNSLDPPLLTRYLRHPSQSWHQAIALRMEVLGCEAQDLY 164
Db 2289 SFTPVNSLDPPLLTRYLRHPSQSWHQAIALRMEVLGCEAQDLY 2332

RESULT 13
US-09-324-867-3
; Sequence 3, Application US/09324867A
; Patent No. 6251632
; GENERAL INFORMATION:
; APPLICANT: Lillcrap, David
; APPLICANT: Cameron, Cherie
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hough, Christine
; STREET: 1669.0010002/JAG/BJD
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/324,867A
; FILING DATE: 1999-06-03
; EARLIER FILING DATE: 09/035,141
; EARLIER APPLICATION NUMBER: 1998-03-059
; EARLIER FILING DATE: 1997-03-06
; EARLIER APPLICATION NUMBER: 60/039,953
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-324-867-3

Query Match 100.0%; Score 871; DB 3; Length 2332;
Best Local Similarity 100.0%; Pred. No. 8.2e-94;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 CDLNSCSMPLGMESKAISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPQVNNPKE 60
Db 2169 CDLNSCSMPLGMESKAISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPQVNNPKE 2228
Qy 61 WLQVDFQKTMKVTGVTTCQVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 2229 WLQVDFQKTMKVTGVTTCQVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 2288
Qy 121 SFTPVVNSLDPPLLTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 164
Db 2289 SFTPVVNSLDPPLLTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 2332

RESULT 14

US-09-315-179-2
; Sequence 2, Application US/09315179
; Patent No. 6376463
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S
; TITLE OF INVENTION: Modified Factor VIII
; FILE REFERENCE: 75-95H
; CURRENT APPLICATION NUMBER: US/09/315,179
; CURRENT FILING DATE: 1999-05-20
; EARLIER APPLICATION NUMBER: U.S. 09/037,601
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: U.S. 08/670,707
; EARLIER FILING DATE: 1996-06-26
; EARLIER APPLICATION NUMBER: PCT/US97/11155
; EARLIER FILING DATE: 1997-06-26
; EARLIER APPLICATION NUMBER: PCT/US94/13200
; EARLIER FILING DATE: 1994-11-15
; EARLIER APPLICATION NUMBER: U.S. 08/212,133
; EARLIER FILING DATE: 1994-03-11
; EARLIER APPLICATION NUMBER: U.S. 07/864,004
; EARLIER FILING DATE: 1992-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-315-179-2

Query Match 100.0%; Score 871; DB 4; Length 2332;
Best Local Similarity 100.0%; Pred. No. 8.2e-94;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CDLNSCSMPLGMESKAISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPQVNNPKE 60
Db 2169 CDLNSCSMPLGMESKAISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPQVNNPKE 2228
Qy 61 WLQVDFQKTMKVTGVTTCQVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 2229 WLQVDFQKTMKVTGVTTCQVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 2288
Qy 121 SFTPVVNSLDPPLLTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 164
Db 2289 SFTPVVNSLDPPLLTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 2332

RESULT 15

US-09-523-656-2
; Sequence 2, Application US/09523656
; Patent No. 6458563
; GENERAL INFORMATION:
; APPLICANT: Lollar S., John
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-95I
; CURRENT APPLICATION NUMBER: US/09/523,656
; CURRENT FILING DATE: 2000-03-10
; EARLIER APPLICATION NUMBER: 09/037,601
; EARLIER FILING DATE: 1998-03-10

; EARLIER APPLICATION NUMBER: 08/670,707
; EARLIER FILING DATE: 1996-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-523-656-2

Query Match 100.0%; Score 871; DB 4; Length 2332;
Best Local Similarity 100.0%; Pred. No. 8.2e-94;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDLNSCSMPLGMESKAISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPQVNNPKE 60
Db 2169 CDLNSCSMPLGMESKAISDAQITASSYFTNMFWATWSPSKARLHLQGRSNARWPQVNNPKE 2228
Qy 61 WLQVDFQKTMKVTGVTTCQVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 120
Db 2229 WLQVDFQKTMKVTGVTTCQVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFGNQD 2288
Qy 121 SFTPVVNSLDPPLLTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 164
Db 2289 SFTPVVNSLDPPLLTRYLRIHPQSWHQAIALRMEVLGCEAQDLY 2332

Search completed: September 29, 2004, 16:41:02
Job time : 34 secs

This Page Blank (uspto)